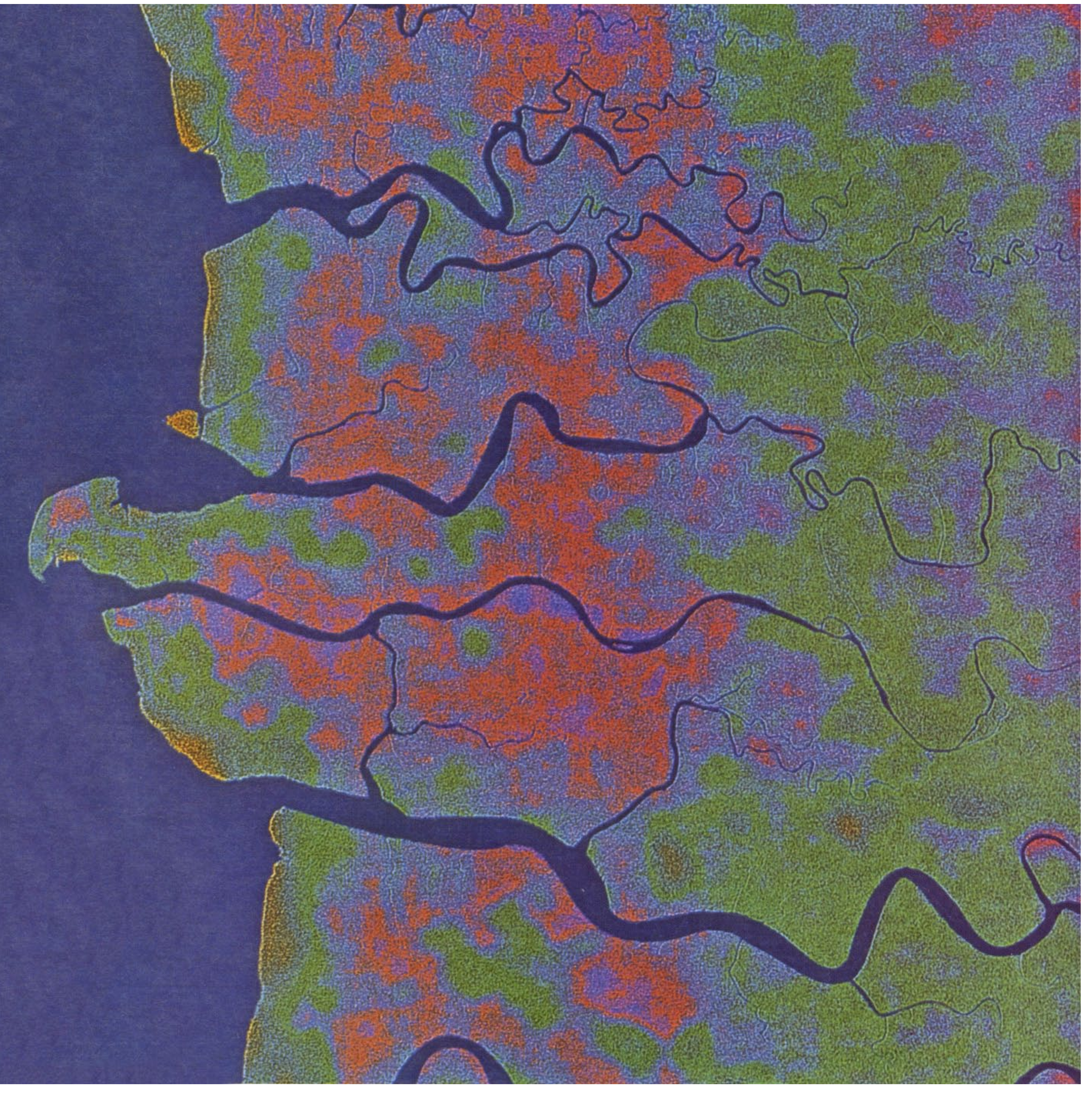
INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN



IMAGENES DE RADAR DE LA TIERRA

Febrero 1983

350 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 77
Febrero 1983

INVESTIGACION Y CIENCIA *Edición en español de* SCIENTIFIC AMERICAN

EDITA Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

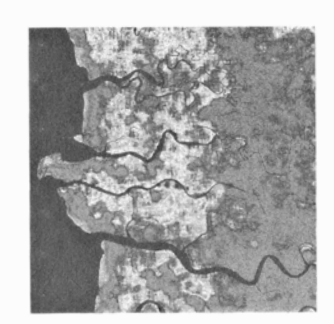
Donato Franco Carrera: *Imágenes de radar de la Tierra desde el espacio*; Mariano Pastor Cruz: *Desarrollo de mapas y bandas en el cerebro*; Luis Bou: *Ordenadores personales*; Manuel Puigcerver: *Muestras de Vía Láctea*; Joandomènec Ros: *Interacción entre la corriente y los organismos sésiles*; Sebastià Xambó Descamps: *A la búsqueda de números primos*; Julio Mangas Manjarrés: *Las lucernas de Cosa*; J. Vilardell: *Taller y laboratorio.*

Ciencia y sociedad:

Juan Modolell

Libros:

Francisco García Valdecasas, Pedro Molera y Juan José Acero



LA PORTADA

La ilustración de la portada es una imagen de radar en falso color de una región costera pantanosa del oeste de Nueva Guinea. Se consiguió con un sistema de radar moderno transportado por la Lanzadera Espacial *Columbia* el año pasado. El radar espaciotransportado, denominado *SIR-A* supuso una nueva clase de sistema de detección remota que amplía el estudio de la superficie terrestre desde naves orbitales a la región de las microondas del espectro electromagnético (véase "Imágenes de radar de la tierra desde el espacio", por Charles Elachi, en este mismo número). Un sistema de imágenes de este tipo se denomina de iluminación lateral y abertura sintética: opera dirigiendo un haz de microondas oblicuamente hacia la superficie de la Tierra y registrando, después, las señales reflejadas y recibidas por la antena del radar mientras la nave se mueve a lo largo de una trayectoria definida; sintetiza luego una imagen de cada punto de la superficie procesando las señales registradas en cada posición orbital donde el punto se sitúa dentro del haz de microondas de iluminación.

Suscripciones:

Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 37

Condiciones de suscripción:

España:
Un año (12 números): 3.850 pesetas

Extranjero:
Un año (12 números): 52 U.S. $

Ejemplar atrasado ordinario:
350 pesetas

Ejemplar atrasado extraordinario:
450 pesetas

Distribución para España

Distribuciones de Enlace, S. A.
Ausias March, 49 - Barcelona-10

Distribución para los restantes países:

Editorial Labor, S. A.
Calabria, 235-239 - Barcelona-29

Publicidad:

Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137 - Madrid-30
Tel. 430 84 81

Cataluña:
Lourdes Padrós
Manuel Girona, 61, 2.° - Barcelona-34
Teléfono 204 45 83

Controlado por O.J.D.

PROCEDENCIA DE LAS ILUSTRACIONES

La fotografía de la portada se publica por cortesía del Jet Propulsion Laboratory, Instituto de Tecnología de California

Página	Fuente
11-19	Miguel Alonso, Julio A. Alonso y José M. López
20-23	Jet Propulsion Laboratory, Instituto de Tecnología de California
24	Ian Worpole
25-28	Jet Propulsion Laboratory, Instituto de Tecnología de California
31	Simon LeVay, Facultad de Medicina de Harvard
32	Martha Constantine-Paton y Margaret I. Law, Univesidad de Princeton
33-38	Tom Prentiss
41	J. Rodón y J. Modolell
46	R. F. Bonifield
47-57	Allen Beechel
58	Hoo-min D. Toong y Amar Gupta, Instituto de
58	Tecnología de Massachusetts
61	Stephen S. Murray, Centro Astrofísico del Observatorio del Harvard College y del Observatorio Astrofísico Smithsoniano
62-64	Gabor Kiss
65	Richard A. Mewaldt, Instituto de Tecnología de California
66-69	Gabor Kiss
71	M. A. R. Koehl, Universidad de California en Berkeley
72-77	Patricia J. Wynne
81	R. F. Bonifield
82-89	Ian Worpole
91-98	Alan D. Iselin
101-107	J. Rodón y M. de Guzmán
111	John H. Lienhard
112-115	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Pedro IV, 160 - Barcelona-5
Fotocromos reproducidos por GINSA, S.A.
Imprime GRAFESA
Gráfica Elzeviriana, S. A.
Napoles, 249 - Tel. 207 40 11
Barcelona-13
Printed in Spain - Impreso en España

Los autores

JULIO A. ALONSO y JOSE M. LOPEZ ("Implantación de iones en metales") pertenecen al departamento de física teórica de la Universidad de Valladolid. Alonso se licenció y doctoró en este centro superior. Pasó luego cuatro años como investigador en la Universidad de Pennsylvania. Sus temas de trabajo actuales son la estructura electrónica de moléculas y sólidos y la termodinámica de aleaciones líquidas, colaborando, en algunos de estos temas, con investigadores de la Universidad de Oxford. López se licenció, por la Universidad de Valladolid, en 1979. Trabaja con Alonso en termodinámica de aleaciones metálicas.

CHARLES ELACHI ("Imágenes de radar de la Tierra desde el espacio") es jefe de equipo en el Jet Propulsion Laboratory del Instituto de Tecnología de California. De su Líbano natal se trasladó a Francia para cursar estudios superiores en la Universidad de Grenoble y en el Instituto Politécnico de esa misma ciudad. Revalidó su título de ingeniero en el Cal Tech, en 1971. Desde entonces trabaja en el Jet Propulsion Laboratory dentro del área de la teoría electromagnética y en la detección de objetos remotos por medio del radar. Dirigió el grupo que diseñó y controló el radar que llevaba la lanzadera espacial *Columbia* en su segundo y tercer vuelo.

MARTHA CONSTANTINE-PATON y MARGARET I. LAW ("Desarrollo de mapas y bandas en el cerebro") investigan en la neurobiología del desarrollo del sistema visual. Constantine-Paton enseña biología en la Universidad de Princeton. Obtenida la licenciatura en la Universidad de Tufts, cursó el doctorado en la de Cornell. "Comencé, escribe, con la neurofisología de los sistemas sensoriales y no ha decaído mi interés por la visión y el procesamiento de la información por parte del sistema auditivo y el somatosensorial. Mi afición por el desarrollo de los sistemas sensoriales tiene su explicación: entiendo que los principios de desarrollo y función de la organización neural tienen que ser similares y que una interpretación mecanicista de cómo se ensambla el cerebro pondría en buen camino la investigación implicada en el funcionamiento del cerebro adulto". Law recibió el doctorado el año pasado en Princeton.

HOO-MIN D. TOONG y AMAR GUPTA ("Ordenadores personales") son ingenieros eléctricos, cuyo trabajo se centra en la interacción entre sistemas de ordenadores y organizaciones humanas. Toong es profesor de dirección de empresas en la Sloan School del Instituto de Tecnología de Massachusetts (MIT) y jefe del laboratorio de sistemas digitales del Centro para Investigación de Sistemas de Información del MIT. Sus grados académicos en ingeniería eléctrica son de ese Instituto, por el que se recibió de doctor en 1974. Gupta cursa estudios de post-doctorado en el MIT. Se formó en la Universidad Politécnica de Kanpur, India, por la que es licenciado en ingeniería eléctrica (1974) y doctor en informática (1980). Desde 1974 trabaja para el gobierno central de la India en la evaluación y adquisición de sistemas de ordenadores. Durante los últimos cuatro años ha dividido su tiempo entre la India y los Estados Unidos.

RICHARD A. MEWALDT, EDWARD C. STONE y MARK E. WIEDENBECK ("Muestras de Vía Láctea") son astrofísicos que han trabajado juntos en el tema de las abundancias relativas de isótopos en los rayos cósmicos. Mewaldt es profesor de investigación del Instituto de Tecnología de California. Licenciado por la Universidad Lawrence en 1965 se doctoró por la de Washington en 1969. En 1971 se trasladó al Instituto de Tecnología de California. Cuenta él mismo que ha empleado datos "de tres vehículos espaciales de la NASA... para investigar la composición isotópica y elemental de núcleos de gran energía acelerados por fuentes galácticas en el Sol y en el espacio interestelar y para estudiar electrones de gran energía procedentes del Sol y de Júpiter". Stone es profesor de física en el Cal Tech, al que se incorporó terminada su carrera en la Universidad de Chicago. Realizó su primer trabajo sobre rayos cósmicos en 1961, con los satélites Discoverer. Desde entonces, "he investigado la composición isotópica y elemental de los rayos cósmicos galácticos y de las partículas procedentes de fulguraciones solares, he estudiado las magnetosferas planetarias y el medio interplanetario y he coordinado las investigaciones de los Voyager referentes al sistema solar exterior". Stone ha desempeñado el cargo de investigador principal para seis vehículos espaciales de la NASA y ha sido premiado con la Medalla de Servicios Distinguidos de la NASA. Wiedenbeck es físico investigador ayudante del Laboratorio de Ciencias Espaciales de la Universidad de California en Berkeley. Licenciado en ciencias en 1972 por la Universidad de Michigan, se doctoró por el Instituto de Tecnología de California en 1978, trasladándose luego a Berkeley. Este año se incorporará a la Universidad de Chicago.

M. A. R. KOEHL ("Interacción entre la corriente y los organismos sésiles") enseña zoología en la Universidad de California en Berkeley. "Empecé en la universidad como estudiante de arte, pero cambié a la biología cuando descubrí que estudiar la forma natural utilizando un enfoque científico me resultaba más gratificante", comenta. Tras licenciarse en arte por el Gettysburg College, en 1970, se doctoró en zoología por la Universidad de Duke, en 1976. Ha pasado por la Universidad de Washington, por la inglesa de York y por la Universidad de Brown.

CARL POMERANCE ("A la búsqueda de números primos") es profesor de matemáticas de la Universidad de Georgia. Comenzó su formación en la de Brown, concluyéndola con el doctorado en la Universidad de Harvad. En 1972 marchó a la de Georgia, donde ha venido enseñando sin interrupción, salvo un curso que dio en la Universidad de Illinois en Urbana-Champaign. Se interesa especialmente por los algoritmos en teoría de números, teoría combinatoria de números, teoría multiplicativa de números y geometría discreta.

CLEO RICKMAN FITCH ("Las lucernas de Cosa") se ha especializado en arquitectura e historia del arte. Cursó estudios de arte, diseño arquitectónico e ingeniería en el George Peabody College para profesores de Nashville. Continuó estudiando arte en el Instituto Watkins y con el pintor Charles Cagle. Sobre su papel en la segunda guerra mundial escribe telegráficamente: "Años de guerra. Trabajé en un despacho de arquitecto en Miami y para el ejército en Fort Oglethorpe, Georgia, como dibujante, haciendo gráficos de máquinas, partes y sistemas para entrenar a las enormes cantidades de mecánicos que se necesitaban en los nuevos ingenios: jeeps, carros, camiones semioruga, tanques". Junto con su marido, el arquitecto e historiador de la arquitectura James Marston Fitch, ha diseñado y construido casas y jardines.

Hace...

José M.ª López Piñero

... cien años

Aureliano Maestre de San Juan publicó en Madrid su *Idea general sobre los organismos*. Se trataba del texto de una conferencia que había pronunciado en un curso de ciencias naturales organizado por el Ateneo Científico y Literario, de Madrid. Fue uno de sus últimos trabajos y en él expuso, desde una perspectiva comparada, una concepción celularista basada principalmente en las ideas de Virchow.

Nacido en 1828 en Granada, Maestre estudió medicina en las Universidades de Granada y Madrid, obteniendo el título de licenciado en 1847. El magisterio de Marcos Viñals lo orientó tempranamente hacia los saberes morfológicos. No obstante, se dedicó al ejercicio profesional, especialmente a la cirugía, hasta que en 1860 ganó por oposición una de las cátedras de anatomía de la Facultad de Medicina de Granada. A partir de entonces, se consagró a la histología, de la que llegó a ser el cultivador más destacado en el mundo académico español con anterioridad a la obra de Santiago Ramón y Cajal. Entre 1863 y 1867 completó su formación con estancias en diferentes laboratorios de Francia, Alemania, Gran Bretaña y los Países Bajos. Su auténtico maestro fue Eloy Carlos Ordóñez, histólogo venezolano discípulo de Charles Robin, que residía en París. Ello explica la gran influencia que las ideas de la escuela histológica francesa ejercieron durante casi una década sobre Maestre, aunque más tarde asumió plenamente los supuestos de la escuela de Rudolf Virchow.

En 1871, Maestre de San Juan se presentó a las oposiciones a la cátedra de segundo año de anatomía de la Facultad de Medicina de Madrid. No las ganó, pero consiguió impresionar a los miembros de mayor altura científica del tribunal, entre los que se encontraba una figura tan significativa como Rafael Martínez Molina. Este prestigio circunstancial, unido al que había alcanzado con sus publicaciones, pesó de modo decisivo en la dotación en dicha Facultad de la primera cátedra española de histología. Nombrado titular de la misma por concurso, Maestre realizó desde ella una labor didáctica ejemplar, no solamente doctrinal sino, sobre todo, práctica. En su laboratorio tomaron contacto con las técnicas histológicas numerosos médicos españoles, entre ellos el propio Cajal, que vio allí las primeras preparaciones microscópicas cuando cursó el doctorado (1877). Por otra parte, se formaron en él sus discípulos Leopoldo López García y Manuel Tapia Serrano. Mientras que el primero continuó la histología académica en la misma línea que Maestre, Tapia trabajó casi exclusivamente en histopatología, especialmente en relación con cuestiones quirúrgicas.

Maestre fue también el fundador de la Sociedad Histológica Española (1874), asociación que logró integrar los esfuerzos de la inmensa mayoría de los cultivadores de la disciplina residentes entonces en Madrid. Figuraron entre sus miembros más activos el catedrático Andrés del Busto y otros médicos no ligados al mundo académico, como el cirujano Federico Rubio, el oftalmólogo Francisco Delgado Jugo y el otorrinolaringólogo Rafael Ariza. También colaboraron personalidades ajenas a la medicina, como el botánico Miguel Colmeiro y el químico Rafael Puerta Ródenas.

Hasta su incorporación a la cátedra de Granada, Maestre publicó trabajos sobre diversos problemas médicos. En uno de ellos, aparecido en 1856, describió por vez primera el síndrome hipogonadismo-anosmia, según han puesto recientemente de relieve García Peris y sus colaboradores. A partir de 1860, su producción escrita, como el resto de su actividad, se centró en la histología. Fue autor de medio centenar de artículos y folletos sobre el tema, así como de un excelente compendio, titulado *Tratado de anatomía general* en su primera edición (1872) y *Tratado elemental de histología* en las dos siguientes (1879 y 1885).

El final de Maestre de San Juan parece simbolizar el carácter heroico de la dedicación científica en la sociedad española contemporánea. En 1888, un accidente en su laboratorio lo dejó ciego. Tras año y medio de duros padecimientos físicos y morales, murió pobre y casi olvidado.

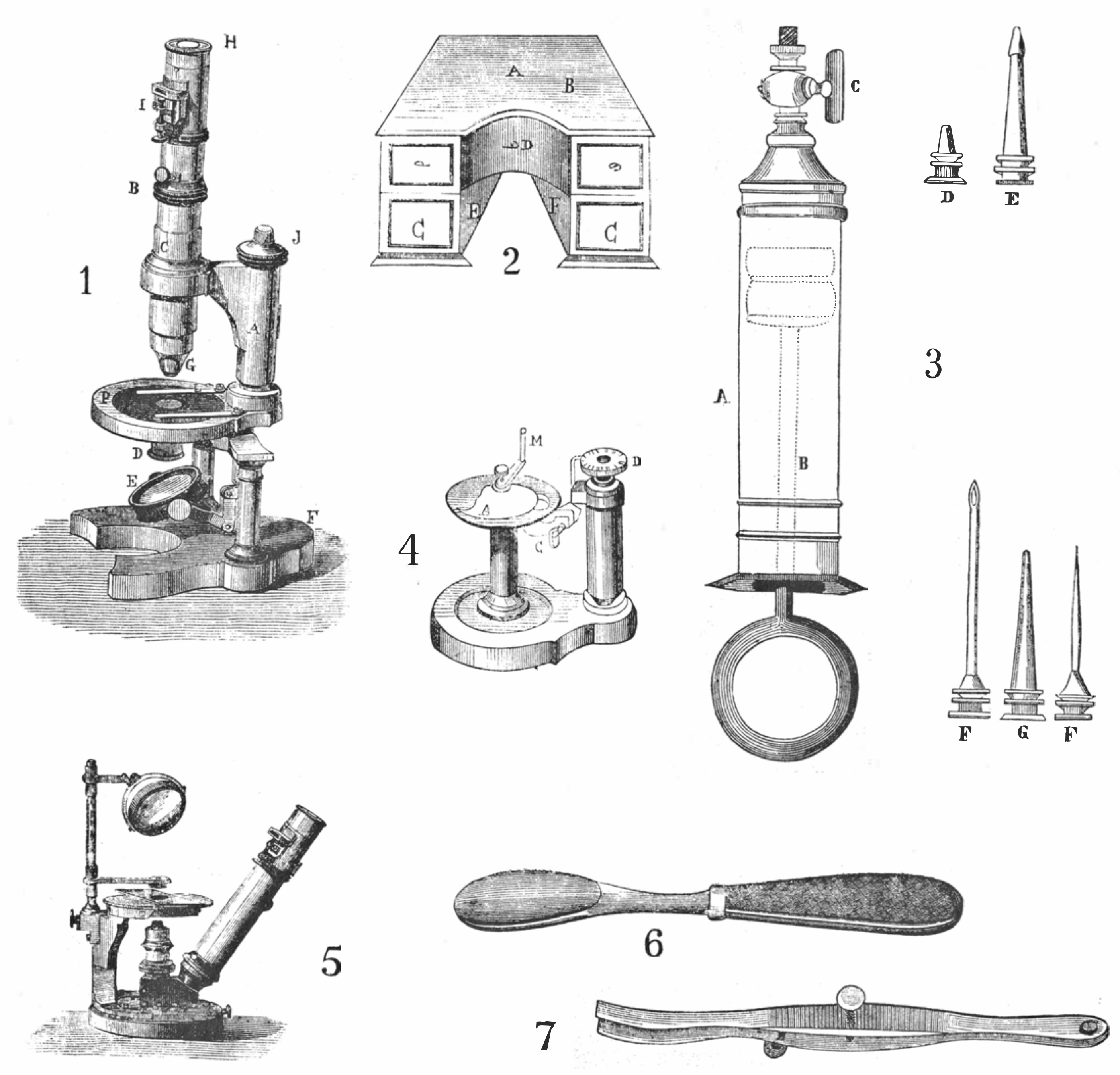
Una torcida mitificación de la figura de Cajal se ha empeñado en presentarle como el primer cultivador en nuestro país de la indagación microscópica. Ello implica desconocer, por una parte, la notable tradición micrográfica española, iniciada a finales del siglo XVII por un investigador de la talla de Crisóstomo Martínez; por otra, el ambiente en el que el propio Cajal se interesó por la histología e inició su formación. Un trabajo de Maestre aparecido en 1875 refleja lo que era la práctica de la histología en España diez años antes de que Cajal comenzase su labor:

"Con satisfacción he podido observar en todas las escuelas médicas y en el mayor número de los hospitales de Francia, Inglaterra, Bélgica, Holanda y Alemania, que he visitado en diversas ocasiones, laboratorios histológicos con cuantas condiciones son de desear; y es muy sensible que en nuestra España, en donde si por fortuna para los progresos de la ciencia médica existe ya una cátedra de Histología normal y patológica en la facultad de Medicina de Madrid, que dispone de un rico laboratorio, y si asimismo el Hospital militar, el de la Princesa, y el Instituto Oftálmico cuentan en la actualidad con su correspondiente laboratorio histológico, servidos todos ellos por un personal entusiasta e inteligente; es de lamentar no se haga esta importante mejora extensiva a los demás de la capital de España, así como a los numerosos hospitales de las diversas provincias...

"Diremos dos palabras respecto al material estrictamente necesario para montar dichos laboratorios, y se verá que su coste es insignificante comparado con los inmensos servicios que a la ciencia reportan. Ante todo es necesario destinar un local a propósito para las observaciones, el cual sea suficientemente capaz: puede constar de dos habitaciones independientes, la una para los trabajos micrográficos y la otra para los químicos, pues no debe olvidarse que las emanaciones de los reactivos alteran fácilmente los microscopios; y a no ser lo primero posible por las condiciones del edificio, debe tener-

se la colección de reactivos en armarios perfectamente cerrados para evitar los perjuicios antes manifestados. La salita de observaciones debe estar iluminada por una sola ventana, pues si hay iluminación lateral no se verá con pureza el objeto; pero si existieran más ventanas, deberán tener cortinas movibles y opacas que las cierren. Utilizando la luz natural, que es la mejor, se preferirá la que proyectan las nubes o un muro próximo y blanco, pues la que se produce por el color azul del cielo no es favorable, así como tampoco la del sol cuando existe movimiento rápido de nubes y agitación de las hojas de los árboles contiguos al local del estudio. Si las observaciones son por la noche, nos serviremos de una lámpara de petróleo, cuyo punto de iluminación debe estar elevado 22 a 40 centímetros por encima de la mesa en que se halla situada, y a una distancia variable entre 25 y 50 centímetros por delante del microscopio según la altura de su foco y la del espejo, y tener una pantalla que impida llegar directamente la luz a los ojos del observador.

"En este local y aproximada a la ventana debe hallarse una mesa para los trabajos microscópicos, que podrá ser de 97 centímetros de altura para trabajar de pie, contando con el desarrollo de tubos del microscopio vertical, formada por un tablero de 125 centímetros de largo y 82 de ancho para que se pueda funcionar con desahogo, de color mate oscuro para que no refleje la luz y fatigue la retina del observador; que tenga a la derecha del mismo tablero y embutida una gruesa lámina de cristal de forma cuadrilátera, de 12 centíme-



Los grabados reproducen otros publicados en la primera edición del tratado de histología de Maestre de San Juan (1872), salvo el segundo, una mesa para trabajos microscópicos, diseñada por el propio Maestre en 1872. Siguiendo el orden numérico: **1**, *microscopio mediano de Nachet;* **2**, *la mesa citada;* **3**, *jeringa y cánulas de Ordóñez;* **4**, *microtomo Nachet;* **5**, *microscopio químico de L. Smith;* **6**, *cuchillo de Strauss y* **7**, *cuchillo doble de Valentin.*

tros cúbicos de extensión, y por debajo de la que se colocarán cuatro fajas de papel de distinto color (negro, blanco, rojo y verde) que servirá para dar cortes a los tejidos, distinguiéndose sus contornos; tendrá dicha mesa dos cajones laterales para encerrar los microscopios en sus cajas, y multitud de instrumentos y preparados, pies gruesos para que den fijeza y estabilidad, los cuales estarán enlazados entre sí los de la derecha y los de la izquierda y los unos a los otros con un travesaño a una conveniente altura para que el observador apoye sus pies cuando trabaje sentado en cuyo caso necesitará una banqueta trípode alta, y de asiento que, aunque relleno, sea bastante duro para que no ceda al peso del cuerpo. Pero si se desea tener una mesa que llene todas las condiciones, puede construirse según el modelo de la que se hizo para el laboratorio histológico de la Facultad de Medicina de Granada, y cuya descripción se encontrará desde la página 48 a 51 de mi tratado de Anatomía general, publicado en Madrid en 1872.

"Los instrumentos absolutamente necesarios los constituyen un microscopio disección Werick, otro de observación del mismo autor, o el modelo mediano Nachet, que tenga los accesorios absolutamente indispensables, como son los micrómetros ocular y objetivo y la cámara lúcida, y si se quiere, además los objetivos de inmersión y de corrección y el aparato de polarizar la luz; un microscopio químico de L. Smith, construido por Nachet; una caja Lebert para preparaciones microscópicas, en la que se encuentran los principales instrumentos para este género de operaciones, como neurotomos, pinzas, tijeras, agujas, el cuchillo doble de Valentin, el microtomo, y el cuchillo de Strauss, etc.; el microtomo de Ranvier, navajas barberas perfectamente vaciadas, sierra pelo de relojero en su arco, piedras de Levante y de América para desgastar laminillas duras, la jeringa de inyección de C. Robin último modelo, o la de Ranvier construida por Mathieu; la prensa de Arturo Chevalier, el aparato de Hett, o el de Cornu para fijar el mastic en las preparaciones, y en número considerable cristales portaobjetos entre los que los habrá con células formadas por cristal, gutapercha, cautchout, o blanco de Ziegler, etc., así como también muchos cubre-objetos por la facilidad con que se rompen en la práctica de las preparaciones.

"Además, debe tenerse en frascos perfectamente tapados, ora líquidos inofensivos como el iodoserum (sin perjuicio de utilizar oportunamente el suero sanguíneo privado de glóbulos por la coagulación de la fibrina, el humor acuoso reciente, la serosidad cefalorraquídea, etc.); ya para endurecer las partes blandas como el alcohol a diversos grados de concentración, y las soluciones de ácido crómico, o las de bicromato potásico; o bien reactivos aisladores, de los cuales den los unos gran transparencia haciendo más visibles los elementos, como ocurre con la glicerina, y los otros destruyan ciertas partes respetando otras, como el ácido acético, el amoníaco líquido, la potasa cáustica, el ácido nítrico, el clorohídrico, las mezclas nitro-clorohídricas, el ácido sulfúrico, el pícrico, el éter, el cloroformo, etc.; los reactivos colorantes, como el yodo, carmín, el picro-carminato de amoníaco, el índigo, la hematoxilina, la fuscina o sea el clorhidrato de rosanilina, el nitrato argentino, cloruro de oro, de paladio, el ácido ósmico, etc.; los líquidos conservadores, de los cuales la glicerina forma la base, como los de Paccini, el bálsamo del Canadá, la gelatina glicerinada, etc.; los cementos formados por el betún de Judea, el blanco de Ziegler, etc.; las materias de inyección de las que en las unas la cola constituye la sustancia coagulable asociada a diversos colores, como el bermellón, amarillo de cromo, o el sulfato de barita; y en las otras, que se titulan frías y que son mejores, las forman principalmente las de azul de Richardson, de carmín de Beale, o blancas de Frey. Se tendrá además de lo dicho algunas pipetas, agitadores y un mortero de cristal, pequeñas cápsulas de porcelana, embudos de cristal con su sostén correspondiente, papel de filtro, lámpara de alcohol, una probeta graduada, varias copas de cristal y una de ellas también graduada, una bureta de Mohr, una pequeña caja de reactivos como las usadas en química analítica sin perjuicio de los reactivos antes indicados, una balanza, frascos de cristal de capacidad varia, de boca ancha y con tapón esmerilado, la mesa metálica trípode de Chevalier; láminas de corcho, de cautchout y de guta-percha, pinceles de pelo de marta y de ardilla, una estufa para secar, lápices de grafito y de color para tomar nota de las preparaciones, valiéndose de la cámara lúcida papel Bristol, cajas para colocar los cristales con su correspondiente preparación, etc., serán entre otros los más indispensables objetos para un regular laboratorio histológico o sea micrográfico, histoquímico y de inyecciones finas, que debe poseer todo hospital importante".

Implantación de iones en metales

Bombardeando la superficie de un metal con iones acelerados de un segundo elemento se obtienen nuevas aleaciones exóticas y de gran interés tecnológico, que no pueden prepararse a través de los procedimientos químicos normales

Julio A. Alonso y José M. López

La forma más estable de la mayoría de los sólidos es la de una red cristalina regular. A ésta se llega por repetición, teóricamente infinita, de la celda unidad, vale decir, de la disposición ordenada de los átomos en el espacio. Una red cristalina de tipo cúbico simple sería la constituida por repetición, en las tres dimensiones del espacio, de un cubo en cada uno de cuyos vértices habría un átomo. Esta regularidad de los sólidos permite explicar muchas de sus propiedades físicas.

Cuando se somete un metal a un proceso de calentamiento y alcanza su temperatura de fusión, la red cristalina se desmorona. ¿Por qué? Debido a las fortísimas vibraciones atómicas. Caída que arrastra dos propiedades: la periodicidad y la simetría. El metal pasa entonces de sólido a líquido. Si invertimos el proceso, tornaremos a la fase sólida con la misma red cristalina que tenía antes.

Quizá sea conveniente introducir algunas nociones fundamentales relativas a la formación de aleaciones. Supongamos una mezcla líquida de dos metales, *A* y *B*. Enfriemos la mezcla hasta que alcance su punto de solidificación. Podemos preguntarnos si la mezcla solidificada es o no un sólido cristalino y en qué sistema reticular cristaliza en el caso de que fuera un sólido cristalino. Para descifrarlo hemos de descubrir primero dos factores: la identidad química de ambos metales y la velocidad de enfriamiento. Si el sólido tarda en enfriarse, nos hallaremos ante un material cristalino que quizá conste de una sola fase cristalina o tal vez de varias fases cristalinas (un conjunto de cristales de dimensiones microscópicas, microcristales, correspondientes a fases cristalinas diferentes).

Los diagramas de fase ilustran ambas posibilidades. Esos esquemas muestran las zonas de estabilidad de las diferentes fases de equilibrio en un sistema formado por dos metales, en función de la temperatura y las concentraciones. Pensemos, a manera de ejemplo, en el sistema formado por plata y oro (Ag-Au). Ambos metales tienen el mismo tipo de red cristalográfica: cúbica centrada en las caras. Forma dicha red una repetición de celdas cúbicas con átomos en los vértices y en los centros de las caras del cubo. Para cualquier concentración, la fase sólida de equilibrio es una disolución sólida con la misma red cristalográfica, esto es, con los átomos de plata y de oro distribuidos al azar sobre las posiciones atómicas de una red cúbica centrada en las caras.

Pero si consideramos otro sistema, verbigracia el que constituyen la plata y el cobre (Ag-Cu), observaremos unas zonas donde coexisten el sólido y el líquido (zonas S + L). Pero no acaba ahí el diagrama de fases relativo a este caso. Apreciaremos también una disolución sólida sustitucional para bajas concentraciones de plata en cobre y otra para bajas concentraciones de cobre en plata, que denominaremos, respectivamente, β y α. En una disolución sólida sustitucional, la red cristalina es la del disolvente. En ella se han sustituido átomos del disolvente por otros del soluto. Las disoluciones sólidas metálicas suelen ser estables sólo para pequeñas concentraciones del soluto, por la sencilla razón de que, al sustituir un átomo del disolvente por otro del soluto, se crean tensiones elásticas en la red del disolvente en virtud de la diferencia de tamaño entre ambos átomos. Para bajas concentraciones de soluto, la red del disolvente puede soportar esas tensiones elásticas, pero a medida que aumenta la concentración de soluto la red se va debilitando, hasta que se llega a un valor crítico, más allá del cual desaparece la disolución sólida. Para concentraciones intermedias del sistema Ag-Cu, la fase sólida estable consiste en una mezcla de microcristales de las fases alpha y beta.

Suele aceptarse el principio general de que cuanto más similares son las características químicas de dos metales, mayor será la posibilidad de que formen disoluciones sólidas concentradas. Aunque la solubilidad mutua de casi todos los pares de metales se ha estudiado en el laboratorio, falta todavía un marco teórico que interprete y dé coherencia teórica a lo observado. Diríase que los parámetros atómicos más adecuados para describir la similitud química entre dos metales son: el tamaño atómico, la electronegatividad (potencia de un átomo para atraer electrones cuando se encuentra en una molécula o un compuesto sólido) y la densidad de electrones en la zona más externa del átomo.

¿Qué ocurre, por último, en un tercer sistema, el formado por el par cobre-oro (Cu-Au)? El diagrama de fases, donde intervienen los parámetros mencionados de temperatura y concentración, nos descubre que la fase sólida de equilibrio a temperaturas moderadamente altas es una disolución sólida para cualquier concentración. A temperaturas bajas, la disolución sólo persiste para pequeñas concentraciones de oro en cobre y pequeñas concentraciones de cobre en oro. Para ciertas concentraciones privilegiadas, hay compuestos intermetálicos: Cu_3Au, CuAu y $CuAu_3$. Cada uno de estos compuestos es sólo estable en un estrecho rango de concentración alrededor de su composición nominal [*véase la figura 1*].

Asimismo, cada uno de esos tres compuestos posee su propia red cristalográfica, distinta de la que caracteriza a los metales puros (oro y cobre). En esa red, los átomos de cobre y oro no se hallan distribuidos al azar, sino que ocupan posiciones no equivalentes y bien definidas. (Si mirásemos instalados en un nudo de la red ocupado por un átomo de cobre, nuestro panorama –el tipo y posición de los átomos inmediatos– variaría del que veríamos instalados en un nudo de la red ocupado por un átomo de oro.) Para composiciones intermedias, una que se encontrara

comprendida entre Cu_3Au y CuAu, se obtendría una mezcla de microcristales, unos en la fase Cu_3Au y otros en la fase CuAu.

En todos los ejemplos ilustrados hasta ahora hemos supuesto que el enfriamiento de la mezcla líquida es lento. Se tiene así tiempo suficiente para que los átomos emigren de unas zonas a otras, se reordenen y den origen a la estructura cristalina de equilibrio. ¿Qué ocurre cuando el enfriamiento es muy rápido? Sabido es que hoy pueden conseguirse velocidades de enfriamiento superiores a los 10^8 grados Kelvin por segundo. En esta situación, los átomos de la mezcla líquida no disponen del tiempo necesario para emigrar a través de la mezcla, por lo que la estructura desordenada y aperiódica de la mezcla líquida se conserva en la fase sólida y resulta una aleación amorfa. El movimiento caótico que existía en el líquido desaparece en la aleación amorfa; ésta, que constituye una fase metaestable, andando el tiempo, evolucionaría hacia una fase cristalina de equilibrio. Pero el tiempo exigido para que esa evolución se lleve a cabo es tan grande que, en la mayoría de los casos, las aleaciones amorfas pueden considerarse estables para usos prácticos.

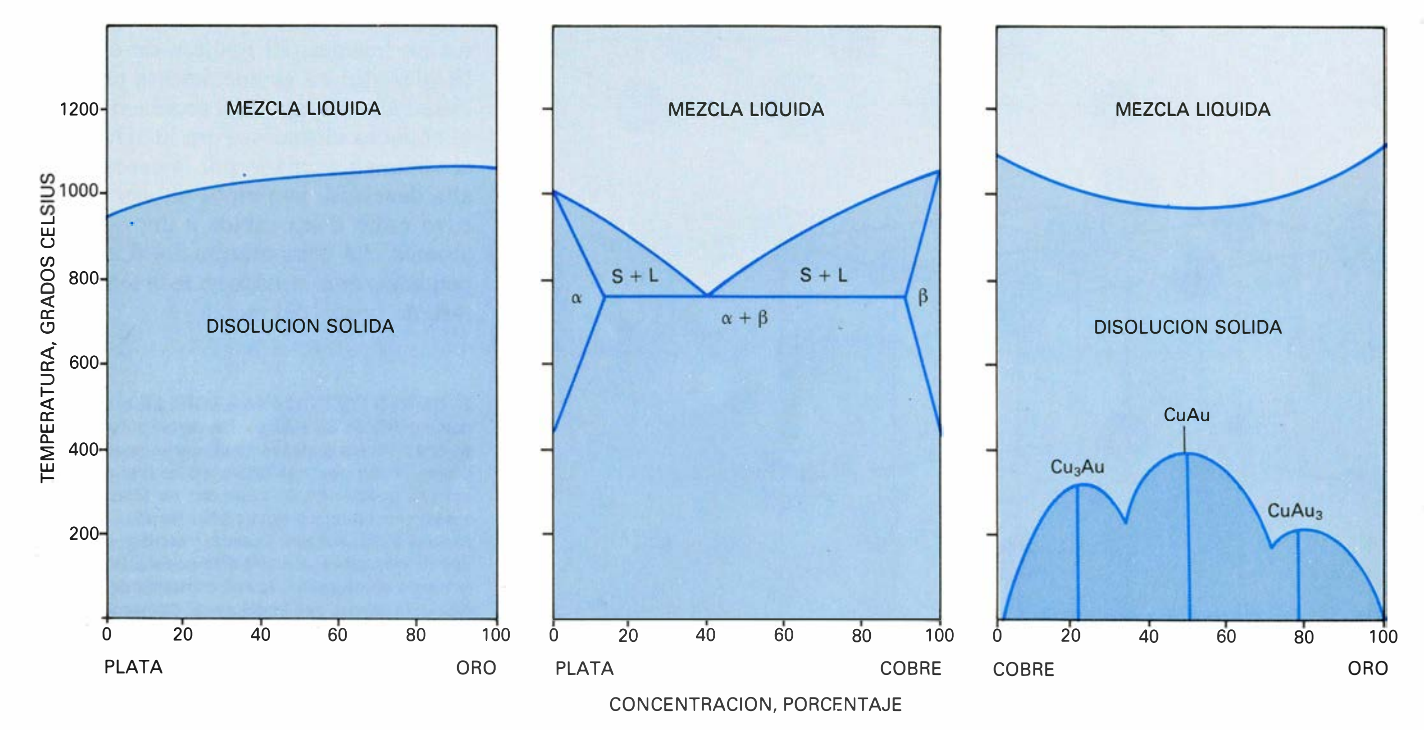
Para conseguir las grandes velocidades de enfriamiento que requieren las aleaciones amorfas, se sigue la técnica de atomizar, en pequeñas partículas, la aleación líquida mediante una explosión controlada. El goteo de materia líquida incide entonces sobre una cinta metálica (de un metal que sea buen conductor del calor), que rota con celeridad. Se alcanzan así velocidades de enfriamiento de hasta 10^8 o 10^9 grados Kelvin por segundo.

En el proceso de obtención de aleaciones sólidas considerado hasta ahora había que partir de una mezcla líquida, obtenida fundiendo juntos los dos metales, que se somete luego a enfriamiento. Con el método de implantación de iones, empero, se forman aleaciones controladas y libres de las limitaciones que imponen los procedimientos térmicos.

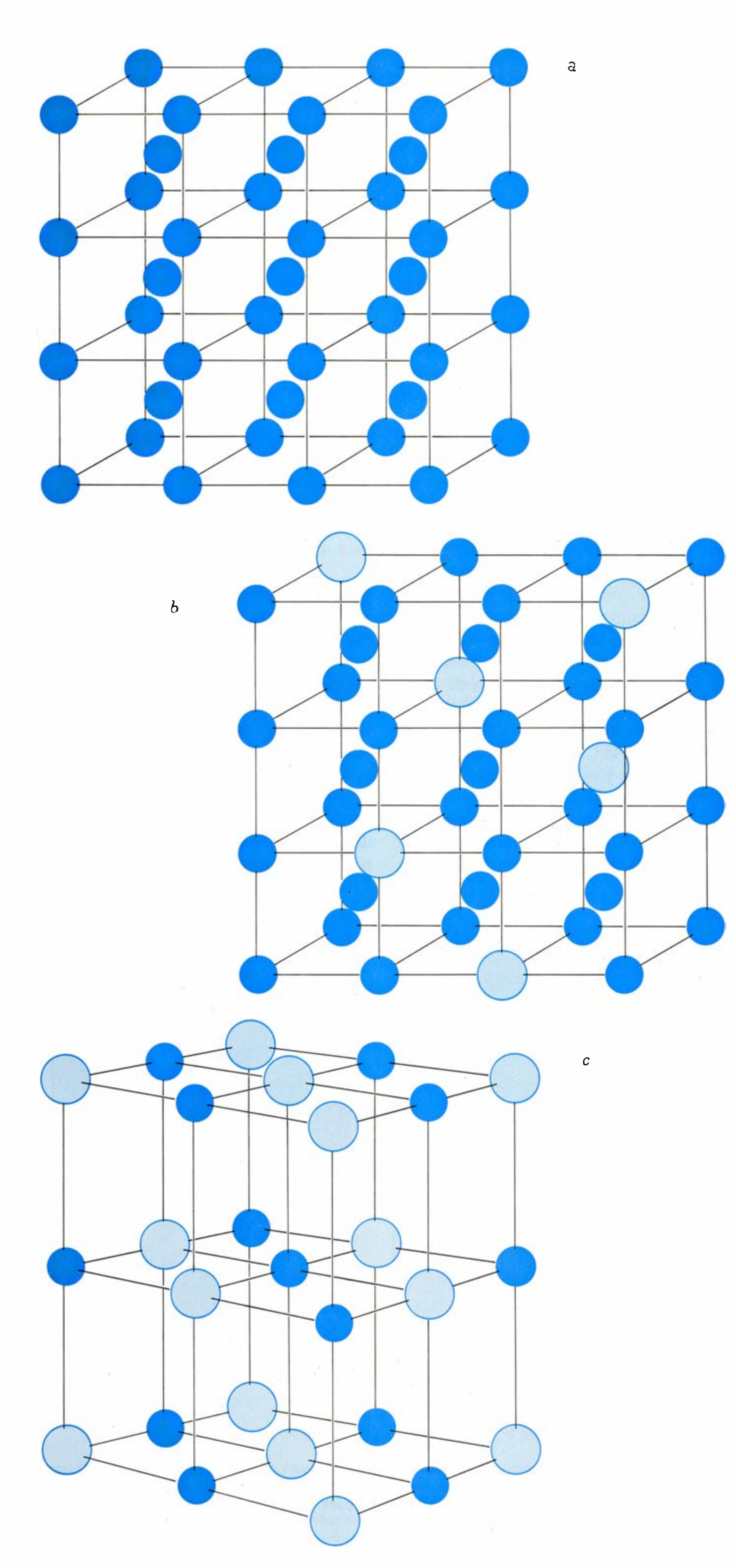
Mediante un acelerador de iones se bombardea la superficie de un metal, llamémosle *A*, con iones acelerados de otro elemento, llamémosle *B*. (Esta técnica se emplea también en semiconductores.) Debido a la gran energía cinética de los iones acelerados (energías de 10^4 o 10^5 electronvolt), éstos penetran directamente en la red cristalina del metal *A*, al margen de cualquier consideración de si *B* es o no soluble en *A*. Obtenemos así una mezcla de composición controlada de átomos de *A* y *B* en la zona superficial del metal. ¿Cómo se controla? Variando el tiempo de implantación y el flujo de iones implantados, donde flujo designa aquí el número de iones que inciden por segundo y por unidad de superficie. La profundidad de la zona implantada, medida perpendicularmente a la superficie expuesta al bombardeo de iones, es de unos 10^{-3} milímetros; lo que significa que las nuevas fases formadas por implantación alteran, sobre todo, las propiedades superficiales del metal.

Mediante el método de implantación de iones se han obtenido disoluciones sólidas sobresaturadas, con una concentración del soluto superior a la que aparece en el diagrama de fases de equilibrio antes descrito, compuestos intermetálicos metaestables, aleaciones amorfas y fases de equilibrio. Ello ha permitido modificar las propiedades superficiales de algunos metales, mejorándose características de los mismos tales como la dureza o la resistencia a la corrosión y al desgaste por rozamiento. De los procesos físicos que se desarrollan durante la implantación de iones y del tipo de aleaciones que se consiguen vamos a ocuparnos de inmediato.

Cuando un ion acelerado penetra en un sólido, aquél cede su energía cinética por medio de dos procesos: primero, por excitación electrónica e ionización de los átomos del blanco; y, en segundo lugar, a través de choques elásticos con los átomos del blanco. Para que ocurra el primer proceso se necesitan energías muy altas (del orden de 10^6 electronvolt). A medida que penetra en el blanco, cada ion incidente desplaza muchos átomos de sus posiciones de equilibrio y éstos, a su vez, des-



1. DIAGRAMA DE FASES de varias aleaciones: plata-oro, plata-cobre y cobre-oro. En todo diagrama de fases se representan las fases estables en función de la concentración y de la temperatura de la aleación. La parte superior es la mezcla líquida y la parte inferior corresponde a las fases sólidas. En la parte *S* + *L* del diagrama Ag-Cu coexisten el sólido y el líquido. En la aleación más sencilla, Ag-Au, hay una disolución sólida continua. El caso opuesto es Ag-Cu, donde las disoluciones sólidas, indicadas por las letras griegas α y β, sólo son estables para una pequeña concentración del elemento minoritario. En la zona de concentraciones intermedias coexisten las fases α y β. Cu-Au presenta una disolución sólida a temperaturas intermedias.



plazan a otros átomos: como si en un castillo de arena introdujéramos un bastón. El proceso se denomina de cascada de colisiones.

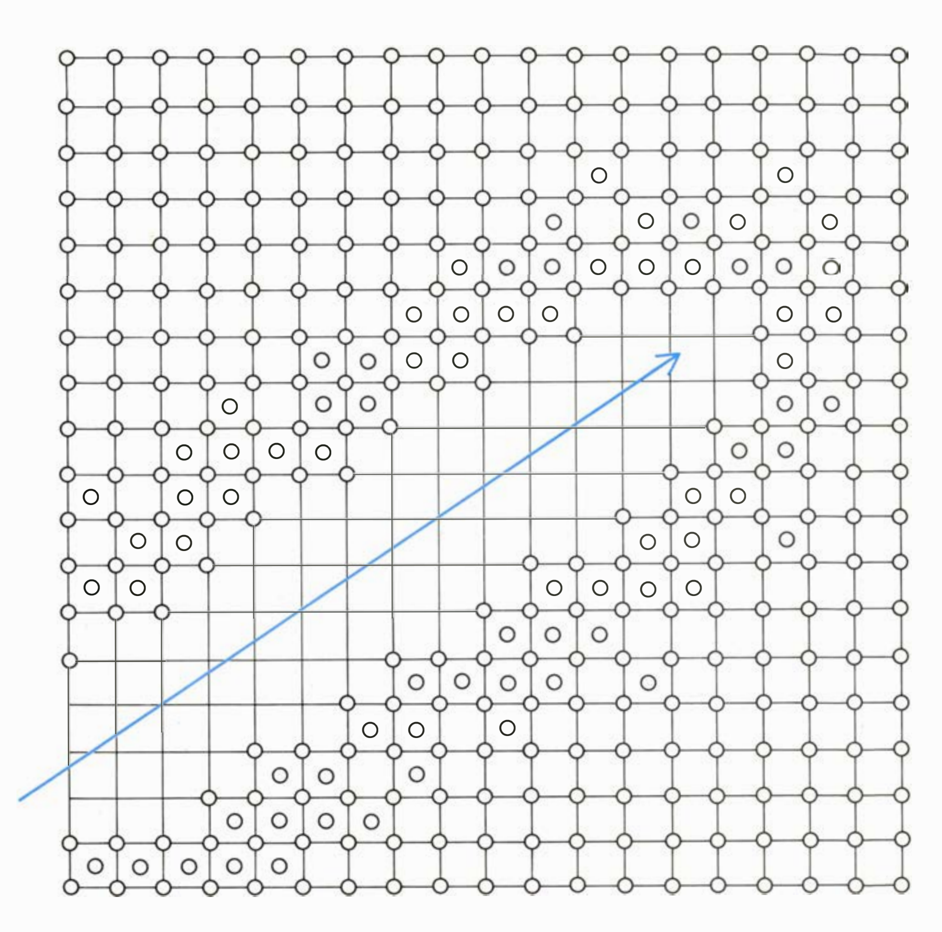
De esa sucesión de colisiones, el tipo más sencillo es el llamado cascada lineal, que ocurre cuando la densidad de colisiones es pequeña, de suerte que cada choque puede considerarse independiente de los demás. En esa situación, el número de átomos desplazados varía linealmente con la energía total cedida por el ion incidente. Si se trata de una densidad alta, entonces el ion cede su energía cinética en un tiempo menor, lo que ocurre cuando la masa del ion incidente y la de los átomos del blanco son altas. Pero ante una densidad de choques muy alta (es decir, cuando la distancia promedio entre dos choques sucesivos llega a ser del orden de la distancia interatómica en el cristal), no nos sirve ya la noción de cascada lineal. Por el contrario, el ion incidente crea una región que, a nivel atómico, aparece violentamente perturbada. Esta región puede abarcar un volumen aproximadamente igual al de la cascada total o bien ocupar ciertos volúmenes localizados (subcascadas) dentro de un volumen mayor, correspondiente a la cascada total. Los volúmenes donde la densidad de choques es alta se denominan cascadas de alta densidad o cascadas de alta energía. En ellas la energía pasa al cristal en un tiempo brevísimo: 10^{-13} o 10^{-12} segundos; aparecen entonces procesos colectivos, donde los choques han perdido su independencia, y efectos no lineales. El número de átomos desplazados no es linealmente proporcional a la energía total cedida mediante choques elásticos. Para imaginarnos el volumen ocupado por la cascada de alta densidad, pensemos en una esfera cuyo radio diera cabida a unos 10.000 átomos. La temperatura local de esa pequeña zona es mayor que la temperatura de fusión del metal.

2. DE RED CRISTALINA REGULAR es la forma más estable de los sólidos. Un metal puro con estructura cúbica centrada en el cuerpo puede considerarse como una repetición, en las tres direcciones del espacio, de un cubo con un átomo en su centro y un átomo en cada vértice (*arriba*). Metales como el sodio, potasio, vanadio y otros poseen este tipo de estructura. En una disolución sólida de un elemento en un metal, la red cristalina de la aleación es la misma que la del metal disolvente en estado puro, pero algunos de los átomos del metal disolvente se han sustituido por átomos de la impureza (*centro*). Por último, la red cristalina de un compuesto intermetálico no guarda relación con las de los metales componentes. Los dos tipos de átomos ocupan posiciones bien caracterizadas. El compuesto representado abajo tiene la misma estructura cristalográfica que la sal común (ClNa). Cada átomo está rodeado por los seis primeros vecinos que son átomos del segundo elemento.

¿Qué cambios sufrirá una red cristalina sometida a la acción de un ion acelerado? Algunos átomos se verán violentamente desplazados de sus posiciones de equilibrio, ocupando posiciones que no son nudos de la red: las llamadas posiciones intersticiales. La energía almacenada en la cascada de alta energía se transmite rápidamente al resto del cristal merced a la conductividad térmica del metal. La temperatura local decae a menos de 500 grados Celsius en un tiempo de 10^{-11} segundos. A lo largo de ese intervalo se reconstruye gran parte de la red ya que muchos de los átomos desplazados vuelven a situarse en los nudos de la misma, no sin dejar un número apreciable de defectos puntuales y defectos extendidos. Defectos puntuales serían átomos en posiciones intersticiales y vacantes, donde vacante significa aquí un nudo de la red cristalina carente del átomo correspondiente; serían defectos extendidos, las dislocaciones, grietas pequeñas y cavidades. Estos defectos extendidos son observables por microscopía electrónica. El enfriamiento tan rápido de la cascada de alta energía es la causa de que en experimentos de implantación de iones se obtengan fases metaestables, por ejemplo disoluciones sólidas sobresaturadas.

Con un abundante flujo de iones implantados, se superponen las cascadas debidas a iones diferentes y el grado de desorden atómico introducido en la red cristalina del metal es muy grande. Este hecho, combinado con el rápido enfriamiento local, explica la frecuente aparición de aleaciones amorfas.

No se conoce todavía bien la naturaleza de la cascada de alta densidad y de sus efectos. Pero algo sabemos. De dos maneras podemos abordar la cascada de alta energía: como una cascada térmica y como una cascada de desplazamientos. Por la primera se entiende un volumen local donde todos los átomos se están moviendo simultáneamente. Este movimiento se puede cuantificar en términos de una energía mínima, la energía de fusión (E_F) del metal. Cuando la densidad de energía elástica (energía cedida por el ion por medio de choques elásticos) es mayor que el calor de fusión del metal, se destruye al instante el orden local de la red cristalina. Al hablar de cascada de desplazamientos queremos señalar un volumen local donde la densidad de defectos (vacantes y átomos intersticiales) producidos por el ion incidente excede un valor crítico, de tal forma que la red se torna incapaz de acomodar estos defectos y sufre una transformación estructural drástica, por ejemplo, el paso al estado amorfo. En este segundo punto de vista, la energía característica es la energía E_d, necesaria para desplazar un átomo de su posición de equilibrio y llevarle a una posición intersticial.



3. EFECTOS PRODUCIDOS en la red cristalina por un ion acelerado que incide sobre el metal. La flecha indica la trayectoria del ion. Los choques de éste con átomos de la red cristalina desplazan a muchos átomos de su posición original, que pasarán a ocupar posiciones intersticiales, es decir, lugares que no son nudos de la red, y dejarán vacantes las posiciones atómicas que ocupaban hasta entonces.

Aunque estos dos enfoques no son equivalentes, ya que E_d es mayor que E_F, guardan una estrecha relación mutua e incluso pueden complementarse para dar una imagen unificada de la formación de una cascada de alta energía. La fase inicial de formación de ésta puede entenderse como una cascada de desplazamientos. Al avanzar el ion incidente, resulta un movimiento neto de átomos del blanco, expulsados hacia fuera de la trayectoria del ion. Este deja así, como rastro, una región rica en átomos que ocupan posiciones intersticiales. Huella que se ha observado directamente por microscopía iónica de campo. La estabilidad del rastro depende de la reacción del cristal a la concentración local de vacantes, ya que la zona colapsada puede reordenarse en gran parte para rebajar las tensiones elásticas introducidas por la creación de defectos.

A medida que el ion avanza, va perdiendo su energía hasta que llega un momento en que no posee la suficiente para provocar más desplazamientos. En este momento, la energía que aún le resta es absorbida por el cristal en forma de vibraciones de la red (calor). Este efecto lo sufren también muchos de los átomos que fueron expulsados de su posición por choques primarios, los debidos al ion incidente, y que, por tanto, tienen la energía aproximada necesaria para crear cascadas térmicas. Como estos átomos son muchos, la zona total ocupada por las cascadas térmicas puede equivaler a la ocupada por la cascada de desplazamiento. La propagación de la cascada térmica tarda unos 10^{-13} segundos; la cascada térmica total y la de desplazamientos se confunden prácticamente.

Acabamos de ver cómo llega la red cristalina a su grado de desorganización local maximo. Pero, ¿cómo se recupera? Antes de contestar a la pregunta y razonar de qué manera desaparecen los defectos creados, vale la pena cerrar el ámbito de las cascadas de alta energía con la descripción de un postrer tipo de cascada de alta energía que aparece en los procesos de implantación de iones en materiales iónicos y en aislantes. En estos casos, y para energías incidentes altas, los choques producen la ionización de varios átomos del blanco muy

próximos entre sí. Debido a la fuerte repulsión coulombiana entre cargas del mismo signo, se produce una explosión coulombiana, en la cual los iones salen despedidos hacia fuera de la zona central, produciendo nuevos choques en su trayectoria, lo que origina una zona rica en vacantes.

El grado de reconstrucción de la red cristalina depende mucho de la temperatura a la que se efectúa el experimento de implantación. El enfriamiento de la zona afectada por una cascada de alta energía es muy rápido y deja abundantes defectos puntuales en la red cristalina: átomos intersticiales y vacantes. Unos y otras pueden moverse posteriormente por efecto térmico, debido a que ningún experimento se realiza a una temperatura de cero absoluto. Para comprender el desarrollo de ese movimiento, atendamos al proceso de reconstrucción de la red cristalina de un metal que acabamos de irradiar con una pequeña dosis de iones a una temperatura cercana al cero absoluto. Supongamos que ahora calentamos lentamente el metal irradiado. Los primeros defectos puntuales que empiezan a moverse son los átomos intersticiales. En su movimiento por la red, algunos átomos de éstos encuentran una vacante y se recombinan con ella restaurando en parte la red cristalina. Los átomos intersticiales que no se han recombinado se agrupan y forman aglomerados de intersticiales.

Si se eleva algo más la temperatura, los aglomerados aumentan de tamaño y se aprecian ya al microscopio electrónico. Algunos aglomerados se encaminan hacia la superficie del metal. A una temperatura superior, las vacantes se tornan móviles. Se debe ello a que, por efecto de la temperatura, los átomos de la red cristalina vibran alrededor de su posición de equilibrio, y la amplitud de esta vibración se alarga con la temperatura. Si se ha producido una vacante en una posición adyacente a un átomo vibrante, éste puede ser atrapado por aquélla cuando la vibración es lo suficientemente intensa, lo que, a su vez, originará una nueva vacante. La vacante se ha movido, diríase, de un nudo de la red a otro nudo. Parte de estas vacantes en movimiento pueden recombinarse con átomos intersticiales. El resto de las vacantes tiende a agruparse formando pequeñas cavidades en el metal, que crecen y se distinguen nítidamente en el microscopio electrónico. Al aumentar aún más la temperatura, las cavidades emigran hasta la superfice del metal o bien se aniquilan con los aglomerados de intersticiales; la restauración de la red cristalina es, entonces, prácticamente total.

La conclusión que se deduce es que la temperatura a la que se efectúa el experimento de implantación condiciona de una manera decisiva el estado final de la red cristalina. En muchos metales, las vacantes tienen una movilidad prácticamente nula a temperatura ambiente. Por tanto, la implantación a temperatura ambiente suele producir una concentración alta de defectos en la red del metal. Por el contrario, si el experimento de implantación se realiza a temperaturas altas, hay una intensa migración atómica durante el tiempo que se prolonga el experimento y el número final de defectos puntuales en el metal será mucho menor.

Hasta ahora no nos hemos preocupado de la posición que ocupa el ion implantado al quedar en reposo al final de su trayectoria por la red del metal. Esta posición final se ha estudiado con detalle en el caso en que la dosis de iones implantados es baja, esto es, cuando el número de iones implantados no llega al 1 por ciento del número de átomos del metal en la zona bombardeada. El resultado es que, al final de su trayectoria, el ion implantado se sitúa, bien en una posición sustitucional, es decir, en un nudo de la red cristalina, o bien en una posición intersticial, con lo que se obtiene una disolución sólida sustitucional o intersticial, respectivamente. Estas disoluciones sólidas suelen ser metaestables: la concentración de soluto trasciende la concentración de equilibrio. Por ejemplo, implantando plata en cobre se han conseguido disoluciones sólidas sustitucionales con una concentración de plata del 17 %; la solubilidad de equilibrio a la temperatura ambiente no llega al 0,1 %.

Fijados el metal *A* y el elemento implantado *B*, este último adquiere una posición única en la red cristalina del

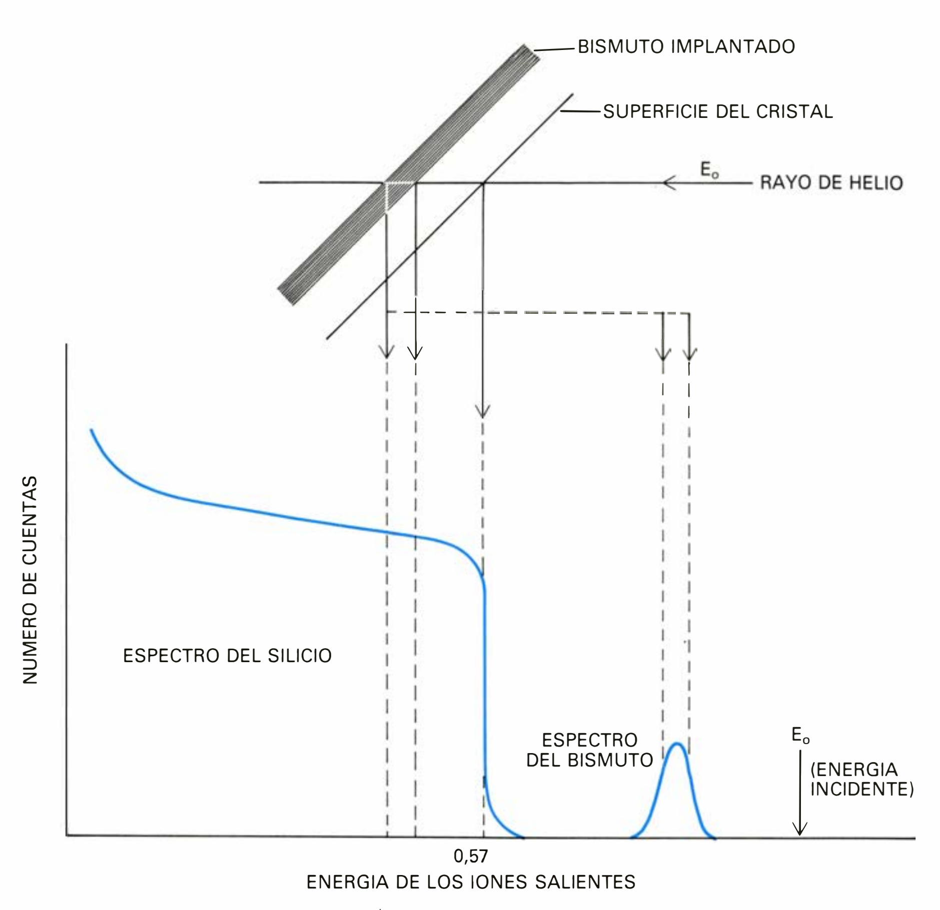
4. AL QUEDAR EN REPOSO, un átomo implantado en un metal adquiere una posición sustitucional o bien una posición intersticial. Un mapa donde las dos coordenadas son la electronegatividad (capacidad de un átomo neutro para atraer nuevos electrones) y la densidad electrónica promedio sirve para separar los elementos intersticiales de los sustitucionales. La figura contiene los resultados de implantación en berilio, obtenidos por Kaufman. Se ha dibujado una elipse que permite separar los dos grupos de impurezas. Podemos predecir qué posición ocuparán los elementos no implantados aún en el berilio.

metal: siempre sustitucional o siempre intersticial. Son clásicos los experimentos realizados por Kaufman y sus colaboradores en los laboratorios Bell a lo largo de los últimos años. Han implantado 28 elementos en berilio y han conseguido observar la posición, sustitucional o intersticial, que adopta cada uno de ellos. Lo crucial aquí se centra en averiguar por qué un elemento siempre adopta posiciones sustitucionales en Be mientras que otro siempre es intersticial. Un interrogante que todavía no ha recibido ninguna respuesta concluyente. Se han avanzado varias explicaciones posibles. La primera es de tipo cinético. Supone este modelo teórico que un átomo implantado termina en una posición sustitucional debido a un choque de reemplazamiento. Como vimos antes, al penetrar el ion en el metal, va perdiendo su energía en los choques sucesivos con los átomos del metal. Se trata de choques de desplazamiento: se aparta un átomo metálico de su posición de equilibrio y se crea una vacante. Cuando el ion se aproxima al final de su trayectoria, ha perdido casi toda su energía en choques anteriores y, al desplazar entonces un átomo de su posición de equilibrio, queda atrapado en la vacante resultante. A ese choque final se le denomina de reemplazamiento.

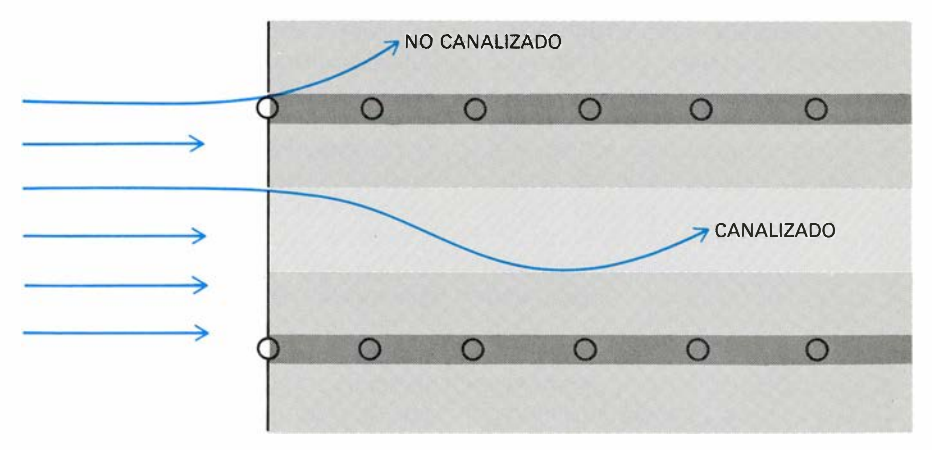
La explicación vale para ciertos metales, aunque no para el caso de implantación en berilio. Puede demostrarse que no hay lugar a choques de reemplazamiento cuando las masas M_A y M_B de los átomos A y B son muy diferentes. La pequeña masa atómica del berilio limita el número de elementos que provoquen choques de reemplazamiento al ser implantados en Be. La experimentación ha puesto de manifiesto, sin embargo, que muchos elementos pesados son sustitucionales en Be. Esto indica que debe haber otras razones para la sustitucionalidad, aparte de los choques de reemplazamiento. Cabe sospechar, en este sentido, que el elemento implantado adquiera la posición sustitucional por recombinación con una vacante de la red cristalina, producida en una anterior colisión de desplazamiento por otro de los átomos implantados. Pero esta hipótesis pierde verosimilitud a temperaturas bajas.

Un segundo tipo de teorías se mueven en un marco termodinámico. Asignan cierta energía E^S a la posición sustitucional y otra energía diferente, E^I, a la posición intersticial; energías que se calculan en función de parámetros atómicos del elemento implantado y del metal. Los parámetros más adecuados parecen ser la electronegatividad, ϕ, y la densidad electrónica promedio, $\bar{n}$, en la zona no muy cercana al núcleo de los átomos. La figura 4 muestra una sistematización de los resultados de implantación en Be. Si se toman como coordenadas ϕ y $\bar{n}$, se observa que los elementos que ocupan posiciones sustitucionales pueden separarse nítidamente del resto en este mapa. En la figura están encerrados dentro de una elipse, lo que refleja cuán idóneos son esos dos parámetros atómicos para describir el fenómeno. Con todo, una teoría completa necesitará probablemente combinar parámetros termodinámicos e ideas cinéticas.

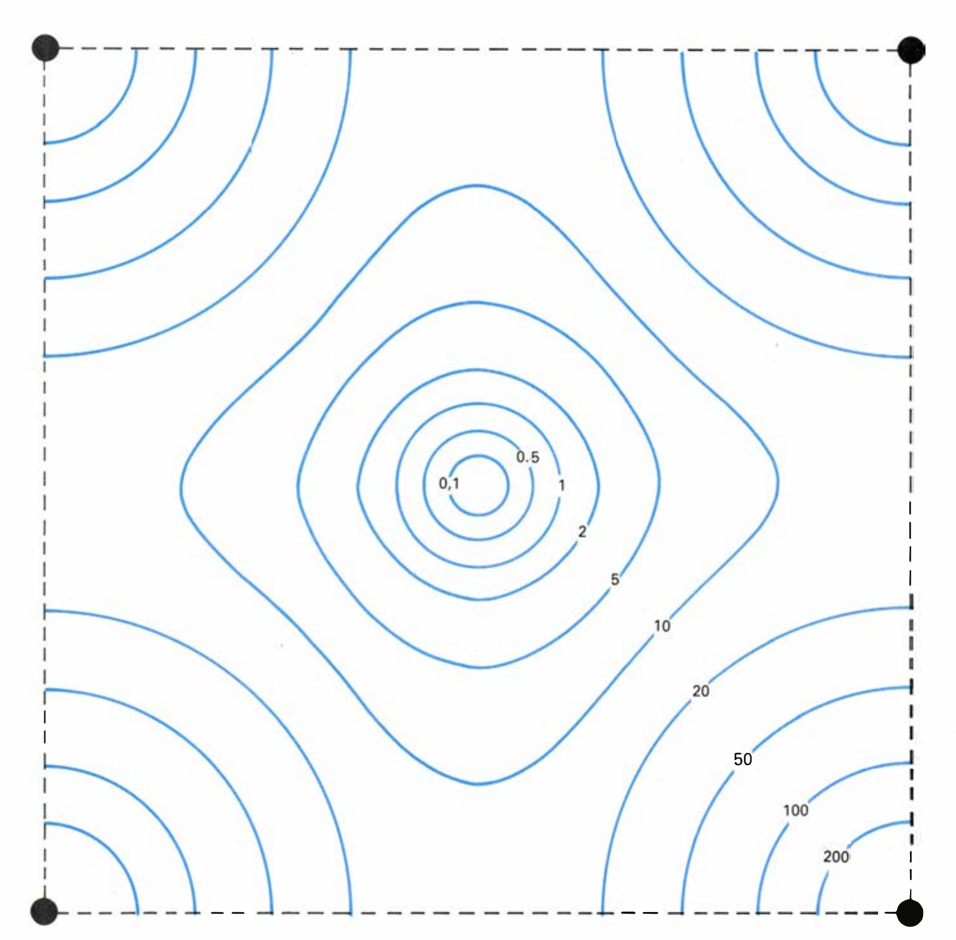
La técnica experimental usada para determinar la posición final, sustitucional o intersticial, de los átomos implan-



5. CUANDO UN RAYO DE IONES He^+ acelerados incide sobre la superficie de un material, penetran en el material, colisionan con los átomos de capas atómicas próximas a la superficie, rebotan y salen del material. Originan así un espectro de dispersión retrógrada ("Backscattering"), en el cual se representa el número de iones desviados en función de su energía. La energía de los iones salientes depende de cuál sea el tipo de átomos en que han rebotado. Por ello, esta técnica sirve para analizar la composición de la zona superficial de una aleación. (Experimento realizado sobre un cristal de Si que ha sido implantado con Bi.)



6. MOVIMIENTO DE UN ION ACELERADO en su penetración en un material sólido a través de la superficie; se ve facilitado si la trayectoria del ion se hace a través de un "canal". Por éste se entiende una zona semivacía del cristal, por ejemplo, la zona entre dos planos atómicos paralelos y consecutivos. El centro del canal es una zona de potencial cercano a cero, limitado por barreras de potencial fuertemente repulsivo (los planos atómicos). Este efecto recibe el nombre de "canalización" y, para que ocurra, el cristal debe estar orientado de forma muy precisa con respecto a la dirección del rayo incidente.



7. CURVAS EQUIPOTENCIALES para un canal de fácil movimiento en un cristal de cobre. La figura representa una sección perpendicular a la dirección del canal. Los círculos negros de cada vértice señalan las posiciones ocupadas por átomos de cobre. Cada curva une puntos con igual potencial. El número que acompaña a cada una indica la magnitud del potencial en esa curva equipotencial. La zona cercana a los átomos de la red es una región de potencial muy repulsivo. Sin embargo, la región central del canal es una zona de potencial próximo a cero, a cuyo través un ion incidente podría moverse fácilmente. El fuerte efecto repulsivo de las paredes del canal tiende a canalizar el movimiento del ion por su zona central.

tados se conoce por canalización ("channelling"). Su alta sensibilidad permite su empleo incluso para el caso de una concentración del elemento implantado muy baja, de hasta el 0,01 por ciento. Se basa en el fenómeno de la difusión hacia atrás "Backscattering". Si se hace incidir un haz de iones de pequeño tamaño y de energía adecuada, He^+ por ejemplo, sobre la superficie de un material cristalino, así un cristal de silicio, parte de los iones He^+ que penetran en el cristal rebotan en las sucesivas capas atómicas que encuentran en su camino, vuelven a salir del cristal y pueden detectarse, dando lugar a un espectro de difusión hacia atrás. En este espectro, donde aparece representado el número de cuentas en función de la energía del ion detectado, puede distinguirse, de la señal debida al rebote en los átomos de impurezas que existan en el cristal, la señal debida al rebote en los átomos del blanco. Esta posibilidad de separar el efecto de los átomos del blanco y el efecto de los átomos de la impureza permite, cuando se combina con el efecto de canalización, determinar la impureza en el cristal.

Si se orienta adecuadamente el cristal respecto del haz incidente de iones He^+ se consigue que los iones que penetran en el cristal lo hagan por canales por donde discurre el ion con facilidad; por ejemplo, el espacio semivacío que media entre dos planos atómicos paralelos consecutivos o el espacio entre dos filas adyacentes de átomos. El canal viene a constituir una zona de potencial cercano a cero, limitada por unas barreras de potencial fuertemente repulsivo. De este potencial repulsivo se responsabilizan las capas electrónicas de los átomos situados en los planos que delimitan el canal. En el centro de este cauce la densidad electrónica es muy pequeña, y, por tanto, es una zona de potencial débil en la que el ion puede moverse sin dificultad. Gracias a esa libertad de acción, el número de iones que sufren difusión hacia atrás disminuye drásticamente, lo que origina una disminución en el número de cuentas en el espectro.

Consideremos el caso en que el cristal contiene una impureza, introducida en un experimento anterior de implantación. Si la impureza ocupa posiciones intersticiales, quedan ciertos canales que pueden detectar la posición de la misma, ya que contienen átomos de dicha impureza. En consecuencia, los iones He^+ que se muevan por esos canales sufrirán difusión hacia atrás al chocar con la impureza, lo que se manifiesta en un aumento de la intensidad en el espectro de difusión retrógrada. Por el contrario, si la impureza ocupa posiciones sustitucionales, el espectro de difusión retrógrada será muy similar al del cristal puro.

De esta forma, estudiando el espectro de iones He^+ canalizados puede determinarse si la impureza ocupa posiciones sustitucionales o intersticiales en el cristal. La técnica faculta incluso para apreciar el tipo de posición intersticial ocupada por las impurezas. En la red cristalográfica del metal berilio, por ejemplo, hay dos tipos de posiciones intersticiales: una con simetría tetraédrica y otra con simetría octaédrica. Puesto que un mismo tipo de posición intersticial está a la vez en varios canales diferentes, variando la orientación del haz de He^+ incidente con respecto del cristal se investigan los diversos canales y se identifica la clase de intersticio ocupado por la impureza.

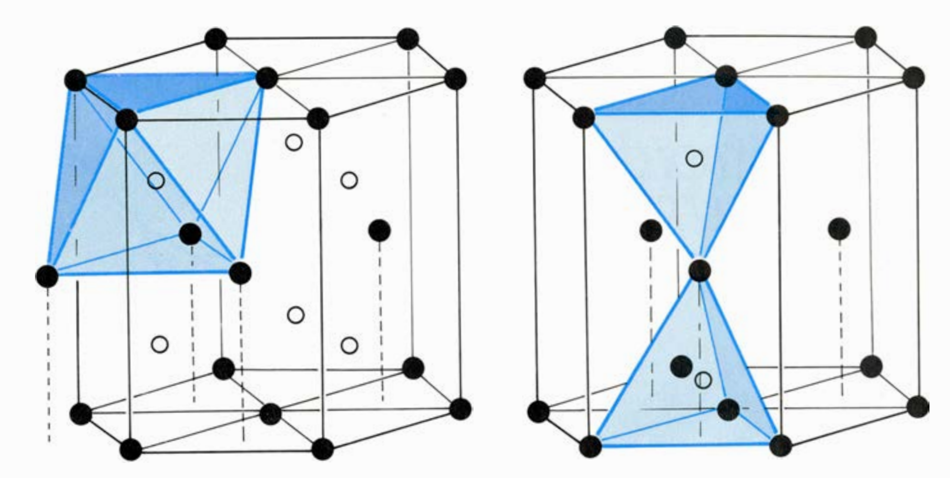
Veamos ahora qué ocurre con la implantación de iones en dosis altas. La diferencia fundamental entre implantación en dosis bajas y altas proviene del hecho siguiente: en la primera, el átomo implantado en un tiempo t_1 no se verá implicado en las reacciones generadas por átomos implantados en un tiempo posterior t_2. Lo que no se cumple para implantación a dosis altas. Atomos implantados en un instante t_1 tienen una probabilidad apreciable de verse implicados en reacciones generadas por átomos implantados en tiempos posteriores. Debido a la alta concentración de átomos implantados, se genera en la red un número muy elevado de vacantes e intersticiales. Las cascadas de alta energía desencadenadas por iones diferentes interaccionan entre sí. Puede crearse tal grado de desorden atómico que la red no sea capaz de restaurarse, sobre todo si la temperatura es baja. Esa es la razón de que aparezcan fases nuevas, algunas de las cuales serán de equilibrio. Otras pueden ser fases metaestables, totalmente nuevas, y éstas merecen abordarse con algún detenimiento.

Tenemos un primer tipo en las disoluciones sólidas sobresaturadas. La red cristalina del metal persiste y se forma una disolución sólida sustitucional o intersticial. La concentración del soluto puede ser varios órdenes de magnitud

mayor que la concentración de equilibrio (recuérdese el ejemplo de implantación de Ag en Cu comentado anteriormente). Una aplicación interesante de estas disoluciones sólidas sobresaturadas es el estudio de la solubilidad de equilibrio como función de la temperatura; otra, la medida del coeficiente de difusión del soluto. El coeficiente de difusión da idea de la velocidad a la que se mueve un átomo de soluto en una aleación. Este movimiento se desarrolla a través de los intersticios de la red cristalina o por sucesivas recombinaciones con vacantes. Para acometer los ensayos pertinentes, se crea una disolución sólida sobresaturada en la zona superficial de un metal, se calienta el sistema y se observa, al microscopio electrónico, la evolución hacia el equilibrio. Ya que la zona superficial es una disolución sobresaturada, el efecto de la temperatura ayudará al exceso de soluto a emigrar hacia zonas del metal alejadas de la superficie. De esta forma puede seguirse, por medio del microscopio electrónico y de la técnica de "canalización", la evolución del perfil de concentración del soluto como función de la distancia a la superficie.

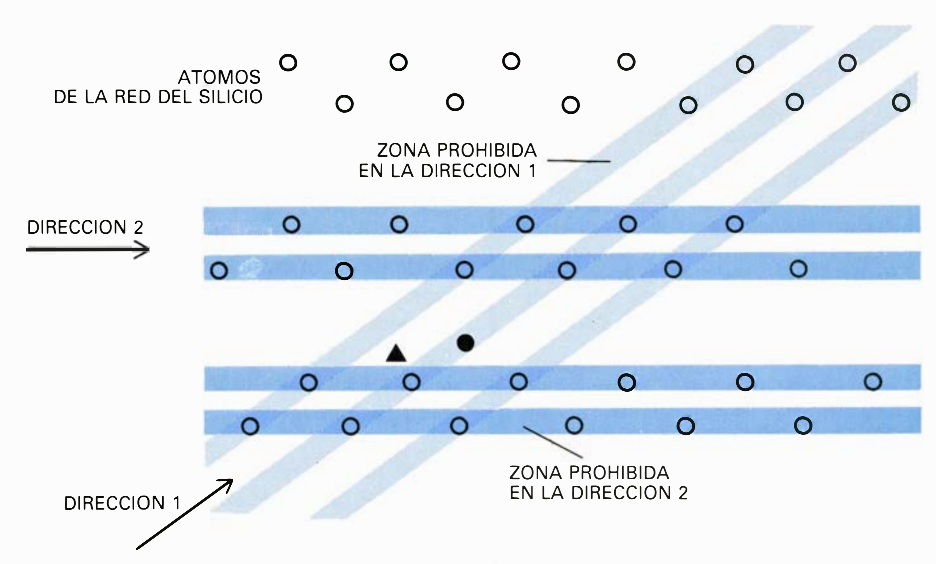
Otro tipo es el formado por compuestos intermetálicos metaestables. Se trata de aleaciones ordenadas de composición fija, que no aparecen en el diagrama de fases de equilibrio. Así, por implantación de platino en silicio se han obtenido los compuestos metaestables $Pt_2 Si_3$ y $Pt_4 Si_9$; por implantación de carbono en hierro, se obtiene el compuesto metaestable $Fe_3 C$.

Nos referiremos, por último, a las aleaciones no cristalinas. Aunque se puede obtener disoluciones sólidas sobresaturadas, el grado de sobresaturación está también limitado. A medida que aumenta la concentración del elemento implantado, se produce una transición de fase sobresaturada a aleación amorfa. Para comprenderlo mejor, pensemos que, al ser alta la concentración del elemento implantado, el desorden atómico introducido en la red del metal adquiere tales proporciones que la red no es ya capaz de restaurarse. Dado el estado actual de nuestros conocimientos, no podemos predecir con certeza qué aleaciones formadas por implantación experimentarán una transición de aleación cristalina a aleación amorfa. Pero sí podemos asegurar lo siguiente: la amorfización ocurre sólo para concentraciones del elemento implantado mayores que el 10 por ciento; el estado amorfo se obtiene de forma sistemática cuando el elemento implantado es un metaloide.

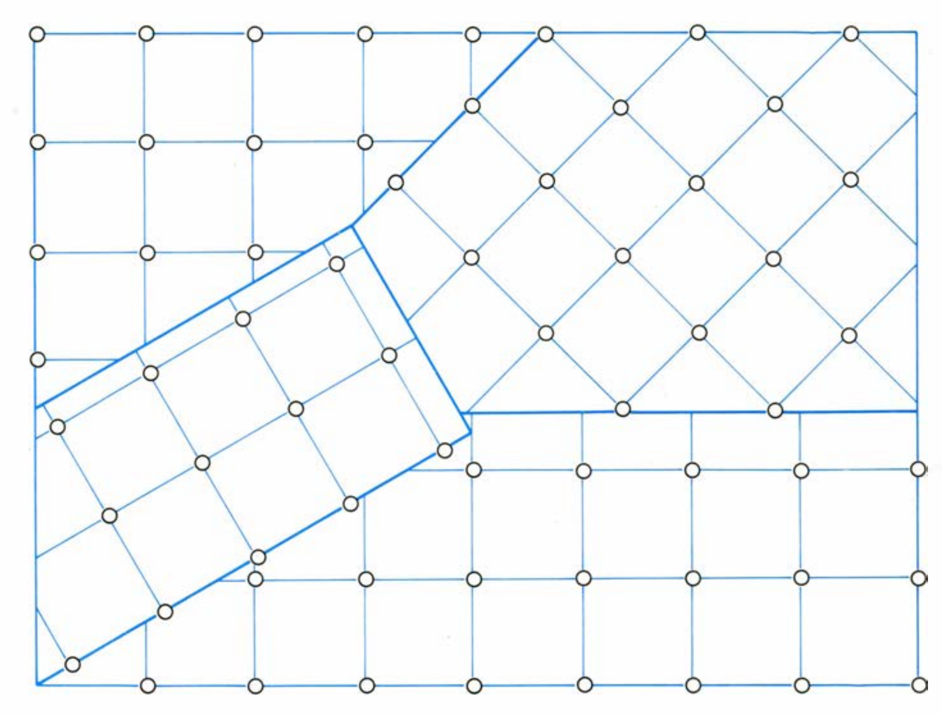


8. INTERSTICIOS en la red cristalina de un metal, donde se amparan las impurezas. Se representa aquí una porción de la red cristalina del metal berilio, de estructura hexagonal compacta. Los átomos de Be están indicados por círculos negros. Hay dos tipos de posiciones intersticiales. Una de ellas tiene simetría octaédrica, es decir, está rodeada por ocho átomos de la red que forman un octaedro. Estos intersticios octaédricos aparecen en la figura de la izquierda como círculos blancos. El octaedro que forman los ochos átomos rodeando a uno de esos círculos blancos se ha dibujado explícitamente. El segundo tipo de posición intersticial tiene simetría tetraédrica, es decir, está rodeado por cuatro átomos de la red (*derecha*).

El interés tecnológico del método de implantación de iones nace de las posibilidades que ofrece de obtener nuevas aleaciones con propiedades prefijadas, capaces de mejorar determinadas características de un material. Los átomos implantados pueden actuar como trampas para capturar impurezas con propiedades indeseables, presentes originalmente en el metal. Ciertos elementos que siempre aparecen como impurezas en los aceros industriales, el antimonio por ejemplo, ejercen un efecto nocivo sobre las propiedades mecánicas de los aceros. El conseguir eliminar esas impurezas durante el proceso de fabricación haría aumentar los costos excesivamente. Aunque la concentración de átomos de Sb es sólo del orden de un átomo en un millón, estas impurezas emigran hacia las superficies de separación de los microcristales del acero. (En realidad, un bloque de un metal o de una aleación no es una red cristalina perfecta, sino un conjunto de microcristales de pequeño tamaño. Dentro de cada microcristal la red cristalina es perfecta, aunque los ejes de los diferentes microcristales no están orientados paralelamente. A las superficies de los microcristales se las llama superficies intergranulares.) Una vez que las impurezas se han concentrado en esas superficies intergranulares, reducen la cohesión intergranular, minando así la resistencia del acero a la



9. PLANO ATOMICO en un cristal de silicio y dos direcciones de posible movimiento en el mismo. Se han dibujado algunos canales prohibidos en la dirección 1 (*las traviesas de la vía*) y algunos canales prohibidos en la dirección 2 (*carriles punteados*). Los átomos de silicio están representados por círculos. El triángulo indica una impureza en una posición intersticial, que podría ser detectada en un experimento de dispersión retrógrada por iones canalizados según las direcciones 1 y 2. El lunar, que representa una impureza en una posición intersticial distinta de la anterior, puede ser detectado por iones canalizados en la dirección 2, pero no por iones canalizados en la dirección 1, ya que está en un canal prohibido para ésta. Tales diferencias permiten determinar la posición ocupada por impurezas implantadas en un metal.



10. CONJUNTO DE MONOCRISTALES, de dimensiones pequeñas, que forma un metal real. Dentro de cada uno, la red cristalina es perfecta. Los microcristales no están orientados paralelamente.

fractura. Sería deseable inmovilizar los átomos de Sb para que no se agrupen en las superficies intergranulares.

En experimentos de implantación de átomos del titanio (Ti) se ha demostrado que el Ti consigue inmovilizar *in situ* a los átomos de antimonio. Tenemos, pues, ahí un medio para aumentar la resistencia de los aceros industriales a la fractura. La mayoría de los metales y aleaciones se transforman en superconductores a temperaturas suficientemente bajas, de unos pocos grados Kelvin. Esto significa que, por debajo de una temperatura T_S, propia de cada material, la resistividad eléctrica cae bruscamente a cero. En aleaciones de paladio con metales nobles (cobre, plata, oro), T_s aumenta al disolver en la aleación pequeñas cantidades de hidrógeno. La concentración óptima de hidrógeno, es decir, aquella para la cual T_s es máxima, está por encima de la que se obtiene a través de procedimientos químicos normales. Dificultad que halla su solución en la implantación directa de hidrógeno en la aleación paladio-metal noble, para conseguir una aleación paladio-metal noble-hidrógeno supersaturada en hidrógeno.

Las aleaciones que se logran por implantación de iones son aleaciones superficiales, es decir, en ese proceso sólo queda modificada una zona de unos 10^{-3} milímetros de profundidad. Significa ello que las propiedades susceptibles de modificación directa se restringirán a las propiedades superficiales del sistema. De donde una aplicación importante la constituirá el aumento de la resistencia de metales y aleaciones a la corrosión por agua o por ácidos. Se sabe que la resistencia de muchos metales a la corrosión aumenta si, en vez del metal puro, se usa una disolución sólida diluida en la que el soluto es un metal noble. Aunque estas aleaciones pueden obtenerse en algunos casos por procedimientos químicos normales, la técnica de implantación permite ahorrar metal noble, ya que sólo se implanta la zona superficial. En otros casos, el elemento noble es insoluble y la técnica de implantación permite obtener una disolución sólida metaestable.

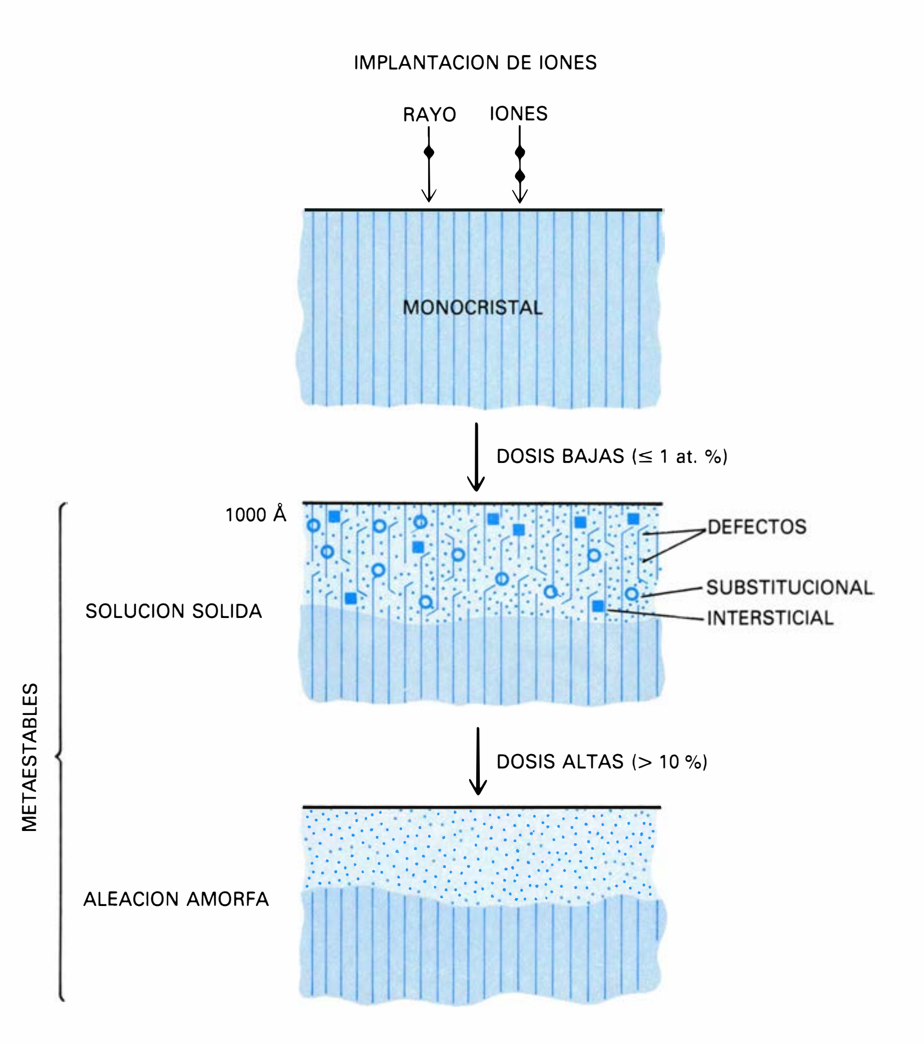
Otra aplicación, ya en uso, es la mejora de propiedades superficiales de piezas mecánicas pequeñas: superficies cortantes o soportes de ejes. La dureza y la resistencia al desgaste de estas piezas puede potenciarse mediante la implantación de algún elemento. Así, el desgaste de un soporte de acero sobre el que gira un eje cilíndrico disminuye a 1/50 del valor original al implantar pequeñas cantidades de nitrógeno en el eje giratorio. Se han aprovechado las vacantes y los átomos intersticiales creados durante la implantación para modificar los mecanismos de transporte atómico y acelerar reacciones (así, la formación de un compuesto intermetálico) en películas metálicas finas.

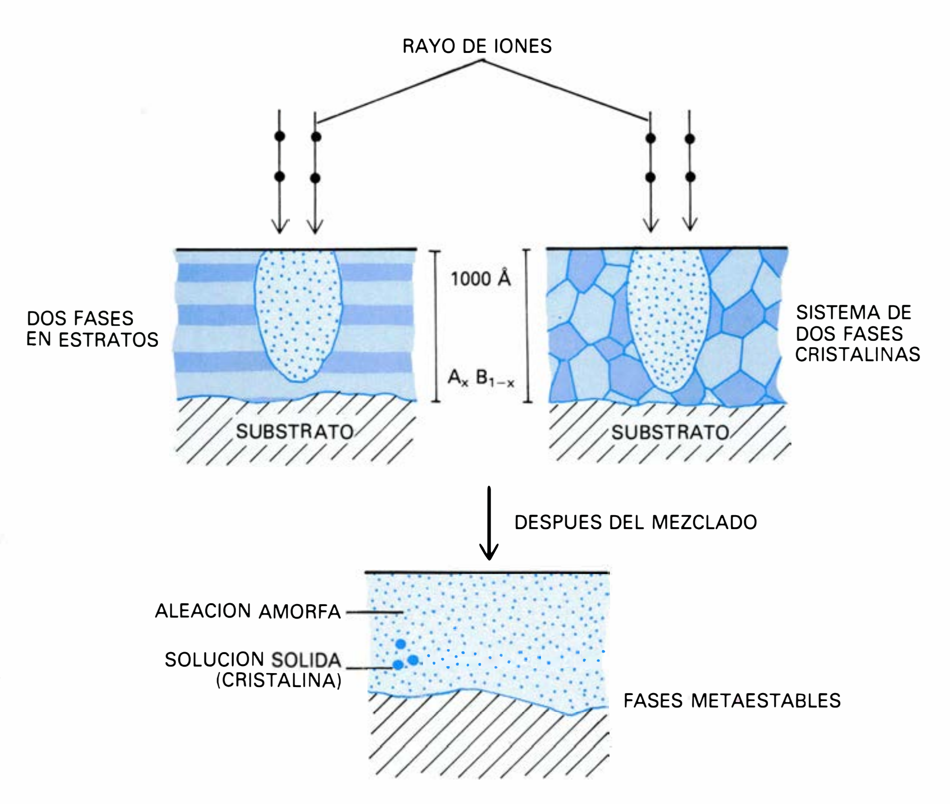
Para alterar las propiedades de un material, la concentración porcentual de iones implantados ha de alcanzar valores apreciables. Dosis altas que no dejan de presentar algunos problemas. En primer lugar, es difícil conseguir

flujos altos de ciertos iones, para los que se precisan largos tiempos de implantación. Por otro lado, un flujo alto e incesante de iones acelerados produce daños en la superficie del material, del que se van desprendiendo las capas atómicas superficiales. Con intención de eliminar estos problemas se ha empezado a usar una nueva técnica, denominada mezcla directa por bombardeo iónico. Se prepara un blanco formando una serie de sucesivas capas finas de dos metales *A* y *B*, dispuestas en la secuencia *ABABAB*. Cada película tiene un espesor de unos 100 angstrom (un angstrom equivale a 10^{-10} metros). Ese emparedado metálico tendrá un espesor total de unos 1000 angstrom. Se bombardea entonces el blanco con iones de un gas noble, argón por ejemplo. El bombardeo produce unos efectos similares a los descritos antes para la implantación directa; pueden obtenerse aleaciones metaestables y también fases de equilibrio.

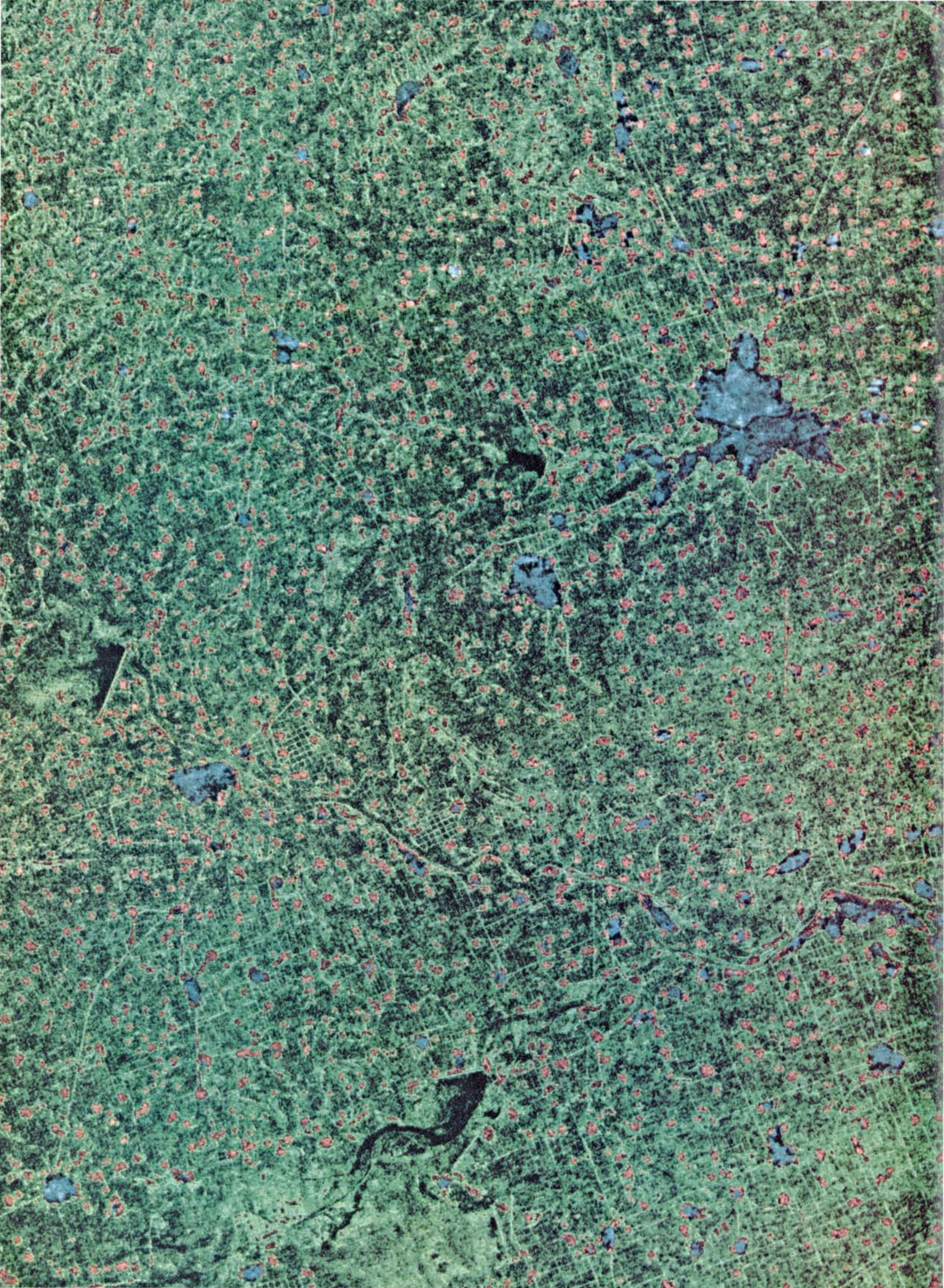
La principal ventaja sobre la implantación directa es que basta una dosis baja de iones de gas noble para producir la mezcla deseada. La pregunta que evidentemente surge es qué ocurre con los iones usados como proyectil. La mayoría de ellos se eliminan a través del siguiente mecanismo: en un primer paso, emigran en buena parte hacia la superficie del blanco, ya que los gases nobles son reacios a formar enlaces. Una vez en la superficie, quedan débilmente ligados a ésta por fuerzas de van der Waals, que ha estimulado la polarización mutua de las nubes electrónicas. Como esto ocurre a la vez que nuevos iones acelerados bombardean la superficie, la energía que estos últimos comunican a la superficie es suficiente para romper los débiles enlaces de tipo van der Waals. Así, los átomos de argón se desprenden del cristal irradiado en forma de gas.

En fase de proyecto aguarda una nueva generación de aceleradores para implantación, mucho más potentes que los actuales. Si hoy las superficies implantadas se cuentan en centímetros cuadrados, con los nuevos aceleradores se podrán contar en metros cuadrados.





11. COMPARACION ENTRE LAS TECNICAS de implantación de iones y de mezcla directa por bombardeo iónico. En la implantación, los iones incidentes actúan a la vez como creadores de desorden atómico y como segundo elemento de la aleación final. En la mezcla directa por bombardeo, los iones incidentes actúan sólo para producir desorden atómico en el blanco, que ya consiste en dos fases que contienen los elementos que se pretende mezclar. En ambas técnicas, el resultado final es una disolución sólida o una aleación amorfa.



Imágenes de radar de la Tierra desde el espacio

Los sistemas de radar en órbita capaces de sintetizar imágenes de la superficie de la Tierra a partir de las microondas reflejadas proporcionan nueva información sobre el relieve

Charles Elachi

La mayoría de las imágenes de la Tierra desde el espacio se han hecho casi siempre a semejanza de las fotografías corrientes: recogiendo, concentrando y registrando la radiación solar reflejada. Otras técnicas de detección remota se basan en la radiación térmica o de microondas emitida por la superficie de la Tierra. El sistema de radar en órbita dirige un haz de microondas oblicuamente hacia la Tierra, detectando después la radiación reflejada y sintetizando una imagen a partir de ella. Los sistemas pasivos de imágenes de alta resolución operan generalmente con luz visible o con radiación en las regiones adyacentes, infrarojo o ultravioleta, del espectro electromagnético. El radar activo espaciotransportado amplía el estudio detallado de la superficie terrestre, desde vehículos en órbita, hasta la región de microondas del espectro, recabando así información sobre las propiedades del relieve que escapaban a toda detección hasta ahora.

Un sistema de radar activo no depende de la luz del Sol; suministra su propia iluminación y funciona, por tanto, a cualquier hora del día o de la noche. Asimismo, estos sistemas pueden obtener imágenes con cualquier clase de tiempo, debido a que las nubes, la niebla y las precipitaciones apenas inciden sobre las microondas. Esta capacidad de hacer imágenes ininterrumpidamente es crucial en la observación de fenómenos dinámicos, tales como las corrientes oceánicas, la migración de hielos flotantes y los cambios de vegetación.

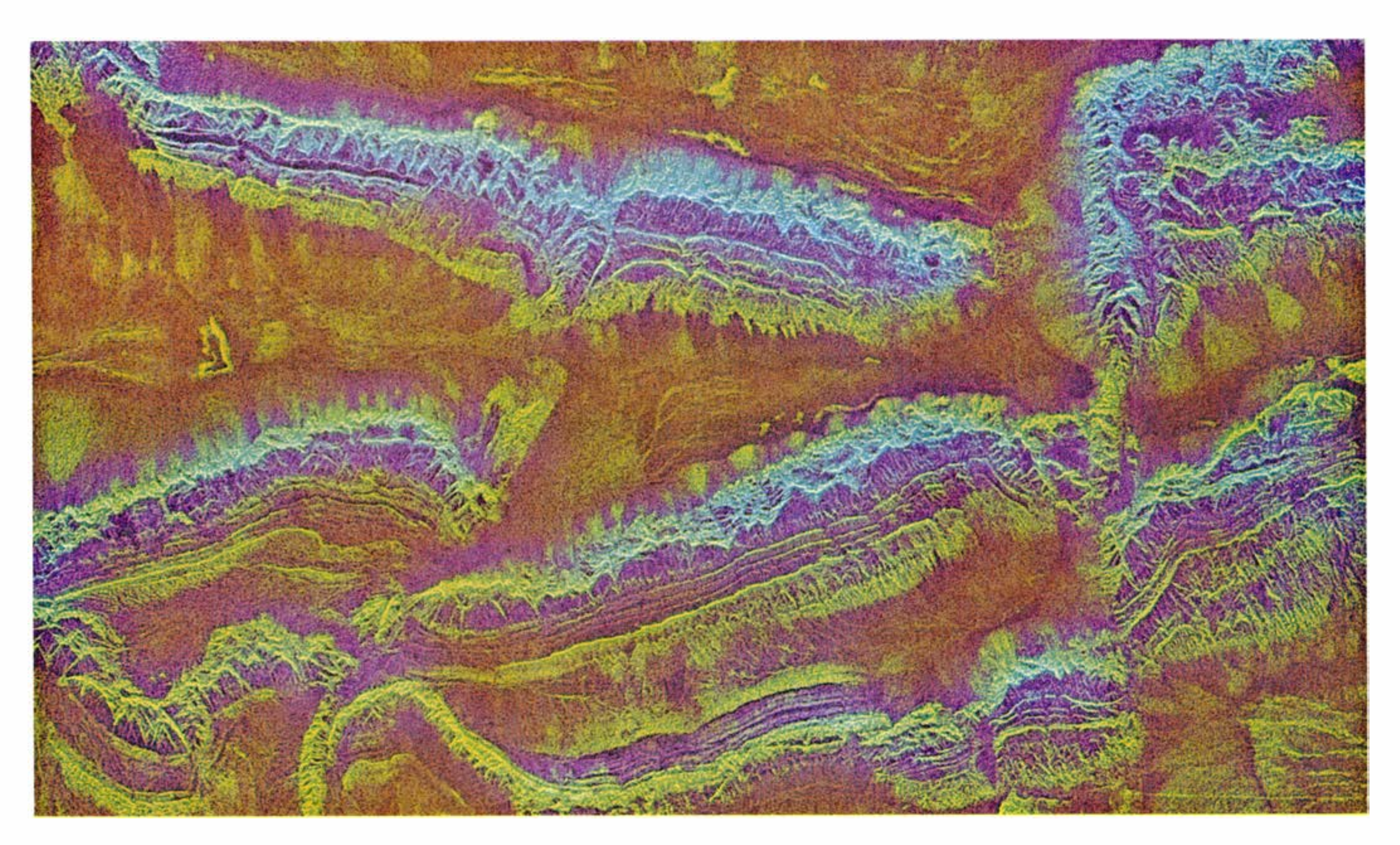
El brillo de un punto dado de la imagen de radar depende de la intensidad de la energía de la microonda reflejada hacia el radar desde el punto correspondiente de la superficie terrestre. La intensidad de este reflejo depende, a su vez, de las propiedades físicas de la superficie (pendiente, rugosidad y capa de vegetación) y de sus propiedades eléctricas (principalmente de su conductividad, relacionada con ciertos factores: porosidad del suelo y contenido de agua, entre otros). En las regiones visible e infrarrojo cercano del espectro, la reflectividad de la Tierra está determinada, sobre todo, por su composición química, mientras que en las longitudes de onda más largas, del infrarrojo térmico, la reflectividad es función principalmente de la capacidad calorífica. En resumen, se necesita una combinación de observaciones de distintos sensores para una descripción completa de la faz de la Tierra.

Hasta ahora han funcionado dos sistemas de imágenes de radar en órbita terrestre. El más reciente iba a bordo de la lanzadera espacial *Columbia* en su segundo vuelo, en noviembre de 1981. Se llamaba *SIR-A*, de *Shuttle Imaging Radar-A* (radar de imágenes de la lanzadera-*A*) y obtuvo imágenes extraordinariamente detalladas de características fisiográficas de todo el globo. Antes, en 1978, voló un radar en el vehículo espacial *Seasat*. Ambos radares se desarrollaron en el Jet Propulsion Laboratory, del Instituto de Tecnología de California.

En un sistema de imágenes óptico o del infrarrojo, la energía emitida o reflejada por una superficie se recoge con una abertura y se concentra sobre un elemento (o un conjunto de elementos) de detección. La resolución angular del sensor está determinada por la razón de la longitud de onda observada al diámetro de la abertura. La resolución espacial es igual a la resolución angular multiplicada por la distancia entre el detector y la superficie. La resolución disminuye con el aumento de la altitud del sensor, con el aumento de la longitud de onda observada y con la reducción del diámetro de la abertura. Como las longitudes de onda de la radiación visible e infrarroja son muy cortas, se pueden conseguir imágenes de alta resolución incluso desde altitudes orbitales de cientos de kilómetros con aberturas de tamaño razonable.

En la región de las microondas del espectro, la longitud de onda es mayor en varios órdenes de magnitud: tanto el radar *SIR-A* como el del *Seasat* tienen una longitud de onda operativa de unos 24 centímetros, alrededor de un millón de veces la de la luz visible. Para que un radar de satélite consiga la misma resolución espacial que otro óptico, desde la misma altura, su abertura debería aumentar en dicho factor. Para obtener una resolución del suelo de 25 metros, por ejemplo, un sistema convencional de imágenes de radar, a una altitud de 250 kilómetros y operando en una longitud de onda de 20 centímetros, necesitaría una abertura de ¡dos kilómetros! Evidentemente, es necesaria otra técnica para generar imágenes de alta resolución con un sensor de radar remoto.

1. PATRON COMPLEJO DE UTILIZACION DE LA TIERRA delineado en esta imagen de radar en falso color de un área rural de la provincia de Hopeh en el nordeste de China. La imagen fue registrada por el sistema del *Radar de la Lanzadera Espacial-A*, o *SIR-A*, desde una altitud orbital de unos 250 kilómetros. El proceso de coloración se hizo en el Jet Propulsion Laboratory, del Instituto de Tecnología de California. Los colores no guardan relación con los naturales: se asignaron basándose únicamente en la intensidad de la radiación de microondas reflejadas, detectada por la antena del radar. Los nódulos de color rojo corresponden a poblaciones. Las áreas verde-oscuras son campos cultivados y las líneas verdiclaras dibujan carreteras y canales de riego. Las formas grises y negras son lagos y embalses, algunos de ellos formados por estancamiento de los cursos de agua. El principal cultivo de la zona es el trigo.



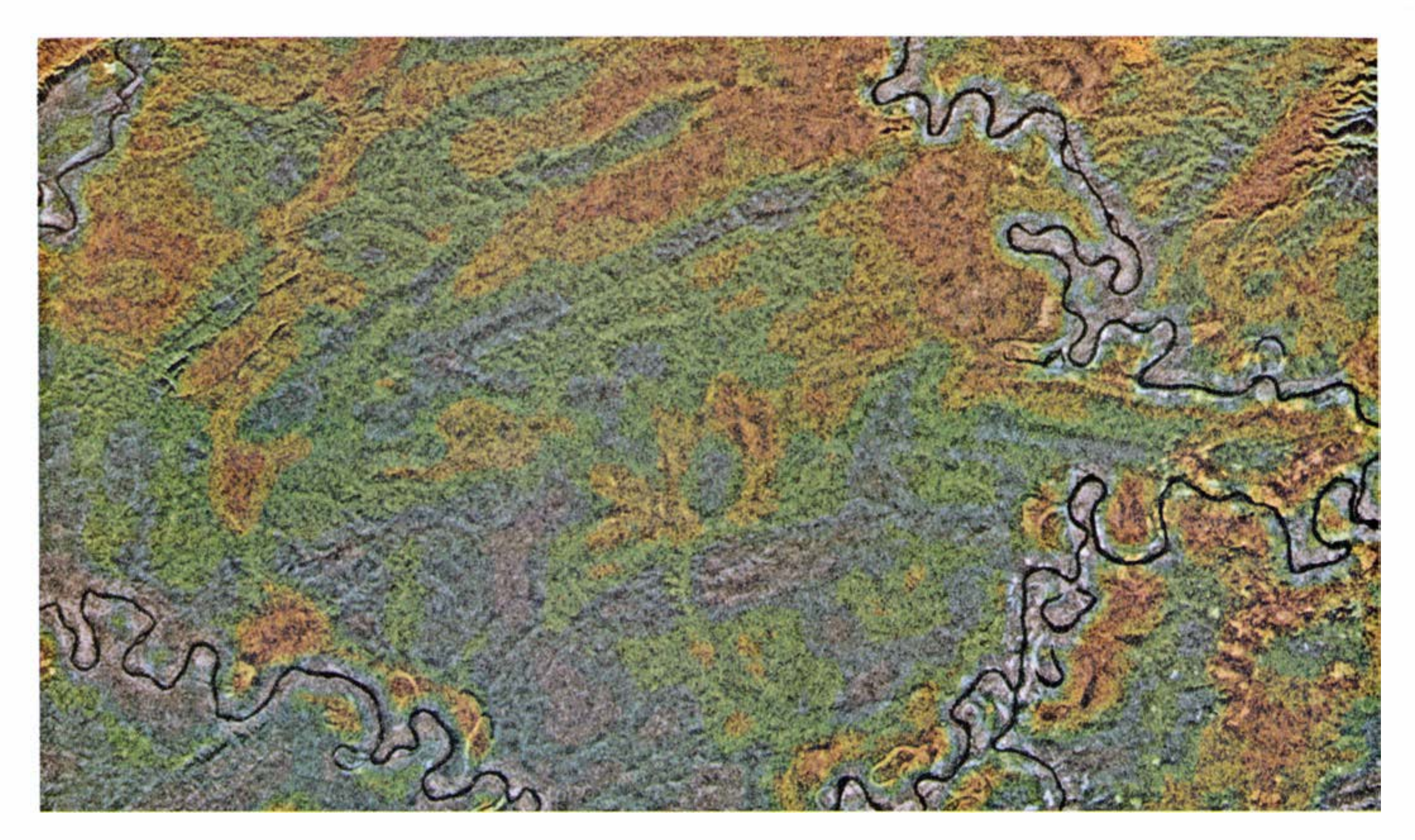
2. GRANDES RASGOS GEOLOGICOS predominan en esta imagen del *SIR-A* de una región árida y montañosa de la provincia de Sinkiang, en el noroeste de China. Las cordilleras se formaron por elevación de rocas sedimentarias; desplazamientos a lo largo de las cordilleras marcan las zonas de fallamiento principal. Los depósitos aluviales formados por erosión de las laderas de las montañas son muy rugosos a la escala de la longitud de onda del radar (24 centímetros) y, por tanto, devuelven una señal reflejada fuerte. Las playas, o valles desiertos entre dos sierras, están constituidas por material más fino y liso, que proporciona un reflejo de radar bastante débil. El plegamiento y la estratificación de las rocas se patentizan en las áreas montañosas.

La solución, desarrollada inicialmente para los sistemas de radar de iluminación lateral aerotransportados (SLAR, de *Side-Looking Airborne Radar*), fue el llamado radar de abertura sintética. En este caso, se hace que una antena de radar relativamente pequeña funcione como otra muy grande (si bien, larga y delgada) aprovechando el movimiento de la antena a lo largo de una trayectoria definida. Una antena grande real (o un conjunto de ellas) envía señales y recoge sus ecos sobre toda la antena simultáneamente. Los ecos de todas las partes de la superficie de la antena se combinan, conforme van llegando, en el sistema de guía de onda del radar antes de que el receptor las detecte. La razón de que la resolución mejore con el aumento de la abertura no se debe simplemente a que una antena mayor recoja señales más fuertes, sino que lo importante es que las ondas son reflejadas desde cada punto de la superficie del blanco a distintas áreas de la antena: al combinarse las ondas reflejadas, interfieren unas con otras según su amplitud y fase. La secuencia de interferencia codifica una información detallada sobre la superficie.

El radar de abertura sintética resuelve también detalles finos al recabar información de las ondas recibidas de un área extensa. La diferencia estriba en que las ondas no se detectan simultáneamente en todas las partes de la antena "sintética", sino que mientras la pequeña antena se mueve a lo largo de su trayecto, se graban las señales recibidas en cada posición para combinarse luego a través de un sistema de proceso de datos. Para establecer una referencia con la que medir la amplitud y la fase de los ecos, éstos se combinan con una señal muy estable conforme los recibe el sensor. En radioastronomía se usa una técnica parecida para formar un interferómetro de muy larga base, combinando las señales de un conjunto de radiotelescopios alejados unos de otros.

En la práctica, una aeronave o un vehículo espacial transportan la antena real, pequeña, a lo largo de la trayectoria; el funcionamiento como antena larga se logra por combinación de los ecos recibidos en diferentes posiciones a lo largo del recorrido. La máxima longitud de la abertura sintética es la distancia a lo largo del trayecto para la cual un blanco dado está dentro del campo de iluminación de la antena. Al aumentar la altitud de la antena, la máxima longitud de la abertura sintética aumenta proporcionalmente, mientras que la resolución espacial alcanzable permanece constante [véase "Radar de imágenes laterales aerotransportado", por Homer Jensen, L. C. Graham, Leonard J. Porcello y Emmett N. Leith; INVESTIGACIÓN Y CIENCIA, diciembre de 1977].

Para el funcionamiento de un sistema de radar de abertura sintética resulta decisivo que las propiedades de las ondas del campo de iluminación puedan ser especificadas y controladas, a diferencia de los de iluminación solar. La iluminación consiste en una serie de impulsos coherentes de microondas en una frecuencia estable y precisa. El retardo entre la transmisión de un impulso y la recepción de su eco proporciona la distancia del blanco. Por consiguiente, las señales de puntos de la superficie que difieren en su distancia pueden separarse según su tiempo de llegada a la antena. Sin embargo, para cualquier posición dada del vehículo espacial, hay muchos puntos con la misma distancia, que forman un círculo centrado en el punto proyección del satélite. Para distinguir las señales de los diferentes puntos con igual distancia se mide el corrimiento Doppler de los ecos: la va-



3. RASGOS GEOLOGICOS MAS FINOS aparecen en esta imagen de radar de una región de gran densidad forestal del sur de México, cerca de la frontera con Guatemala. La imagen se tomó por el sistema de radar transportado por la nave *Seasat* en 1978. Se le aplicaron técnicas especiales de proceso de datos, modulando el color de la imagen en relación con las variaciones espaciales de baja frecuencia (o de gran escala) y el brillo de la imagen en relación con las de alta frecuencia (o de pequeña escala). De esta forma, se pueden resaltar estructuras poco visibles, como las fallas. La confluencia de los ríos Lacantum (*arriba*) y Salinas (*abajo*) cae justo fuera del borde de la imagen, hacia el noroeste (*derecha*). (Cortesía del Jet Propulsion Laboratory.)

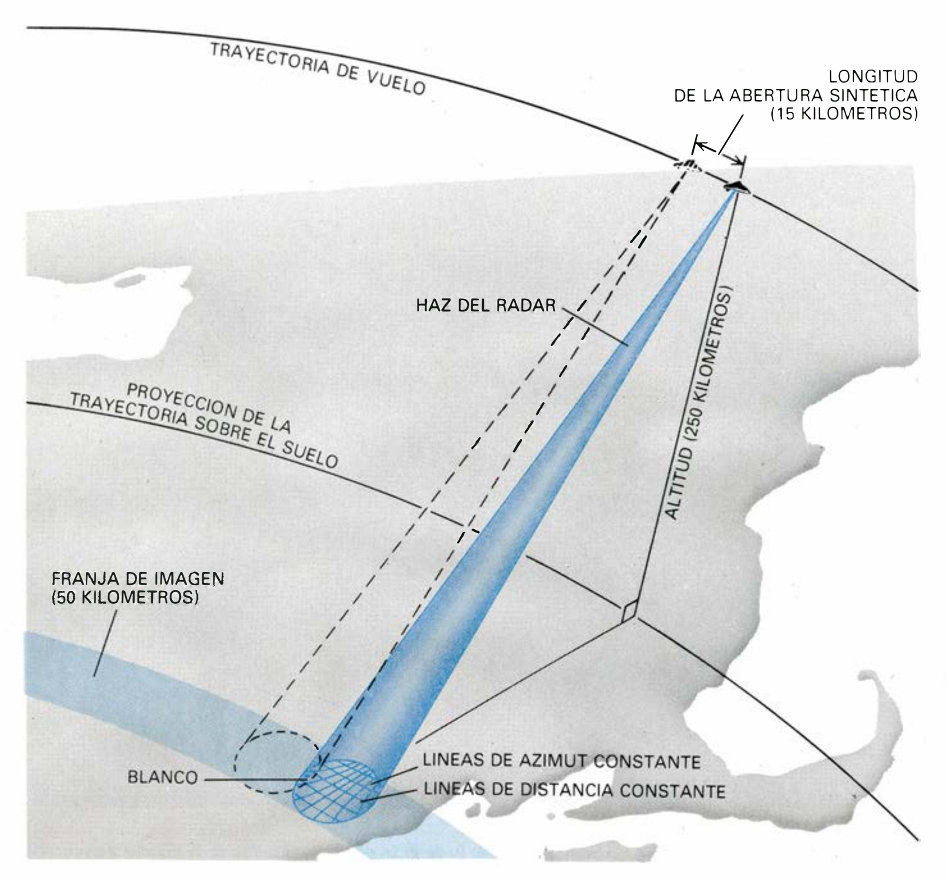
riación de frecuencia producida por el movimiento relativo del vehículo espacial y el blanco. Solamente la componente del movimiento paralela a la dirección del haz de microondas contribuye al corrimiento, por lo que el eco de un punto perpendicular a la trayectoria del vuelo tiene un corrimiento Doppler de cero. Los puntos por delante de la perpendicular se desvían a frecuencias más altas; los de detrás, a frecuencias más bajas. Los datos combinados del retardo y el corrimiento Doppler de los ecos recogidos durante el paso del sensor sobre un área determinada son procesados para aislar la energía devuelta por cada elemento de resolución de la superficie. Así, analizando la información del retardo y el corrimiento Doppler de los ecos se puede cartografiar la energía recibida en una imagen bidimensional que represente, punto por punto, la energía reflejada.

La resolución de un radar de abertura sintética está limitada por la capacidad del sensor para medir con exactitud las pequeñas diferencias en el retardo y el corrimiento Doppler de elementos de superficie adyacentes. La exactitud de tales mediciones no depende de la distancia del sensor a la superficie. Así pues, mientras la señal de retorno esté suficientemente por encima del nivel de ruido, la resolución alcanzable con esta clase de sistema de imágenes resultará esencialmente independiente de su altitud. Tanto el radar *SIR-A* como el del *Seasat* proporcionaron imágenes con una resolución del suelo de 25 metros, en una longitud de onda de 24 centímetros y con una antena de 10 metros de largo. Esta resolución supera en un factor de tres la conseguida por los sensores anteriores del Landsat, que se basaban en la radiación solar reflejada del visible y del infrarrojo cercano.

Para analizar la historia completa del retardo y el corrimiento Doppler de los ecos registrados mientras el sensor estudia un elemento de resolución, se requiere un gran número de cálculos y manipulaciones de datos. En el caso del radar del *Seasat*, hubo que realizar unas 1000 operaciones complejas para sintetizar la abertura larga y determinar el brillo de un solo pixel, o elemento de imagen, de 25 metros de lado. El satélite se movía a unos 7,5 kilómetros por segundo; el sensor cubría una franja del suelo de 100 kilómetros de ancho, con lo que fueron necesarias unas mil millones de operaciones para generar una imagen de alta resolución a partir de la información registrada en un segundo. Además, el procesador de datos debía tener en cuenta los corrimientos Doppler producidos por la rotación de la Tierra y las distorsiones causadas por las pequeñas variaciones de la altitud y la actitud del sensor.

La gran demanda de proceso de datos del radar espaciotransportado de abertura sintética ha supuesto un mayor reto para el desarrollo de tales sistemas. El CRAY-1, uno de los computadores en uso de mayor celeridad, es capaz de ejecutar unos 100 millones de operaciones por segundo, lo que significa un orden de magnitud más bajo que la tasa necesaria para el procesado de datos en tiempo real del radar *SIR-A* o del *Seasat*. Se han desarrollado dos métodos para realizar esta prodigiosa tarea de proceso de datos: uno de ellos óptico y el otro digital.

Al registrar ópticamente los ecos sobre una emulsión fotográfica, el reflejo de un blanco puntual forma una serie de anillos concéntricos característica, similar a la figura de las zonas de Fresnel que se encuentra normalmente en óptica. Una onda óptica plana con este comportamiento puede concentrarse en un punto, que será una imagen del blanco. Este sencillo procedimiento



4. FUNCIONAMIENTO DEL RADAR DE IMAGENES ESPACIOTRANSPORTADO, basado en el principio de abertura sintética como se muestra en este diagrama esquemático. Mientras el sistema de radar de iluminación lateral se desplaza a lo largo de su trayectoria de vuelo, envía un haz de impulsos coherentes de microondas oblicuamente hacia la superficie de la Tierra. Se registran las señales reflejadas recibidas por la antena del radar. Las señales detectadas en varios puntos a lo largo de la trayectoria orbital se combinan posteriormente por un sistema de proceso de datos para formar una imagen bidimensional del área del blanco. En efecto, la antena de radar en movimiento actúa como una antena alargada (o como un extenso conjunto de pequeñas antenas); la máxima longitud de abertura sintética resultante es la distancia según la trayectoria para la que un blanco dado se encuentra dentro del haz de iluminación de la antena. Cada punto de la superficie se caracteriza por dos magnitudes: su distancia, que es función del retardo entre la transmisión de un impulso y la recepción de su eco, y su azimut, que se determina midiendo el cambio de frecuencia producido por el movimiento relativo de la nave y el blanco (el efecto Doppler). La imagen se compone analizando toda la información sobre el retardo y el corrimiento Doppler.

(que puede considerarse como analogía bidimensional de una holografía) constituye el fundamento de todas las técnicas ópticas de proceso de datos de un radar de abertura sintética. Los procesadores ópticos basados en este principio comenzaron a diseñarse en los años cincuenta y principios de los sesenta por miembros de la Universidad de Michigan y de la Goodyear Aerospace Corporation para elaborar datos de radares de abertura sintética aerotransportados. Posteriormente, se adaptaron los sistemas ópticos al procesado de datos de los radares *SIR-A* y del *Seasat* en el Jet Propulsion Laboratory.

La segunda aproximación al problema del proceso de datos depende por completo del computador digital. El gran número de cálculos necesarios ha limitado la capacidad actual de las operaciones, que no intentan seguir la corriente de datos conforme sale del sensor. El procesador digital más rápido disponible para radares de imágenes espaciotransportados tiene una "razón de paso" de 1 a 500; esto es: los datos recogidos en un segundo requieren 500 segundos de procesado. El procesador fue desarrollado en el Jet Propulsion Labortory por un equipo encabezado por Chialin Wu. En Gran Bretaña, Canadá, Alemania y Japón se han construido procesadores más lentos. El Jet Propulsion Laboratory tiene en sus manos el diseño de un procesador que igualaría la cantidad de datos del sensor. Su terminación está prevista para 1985.

Las imágenes de radar de la Tierra desde el espacio encuentran aplicación generalmente en ciertas áreas de geología, planetología, oceanografía y el sector de los recursos renovables. Un campo donde la explotación de los datos de radar está particularmente avanzada es la cartografía del relieve. Muchas características geológicas superficiales o próximas a la superficie, tales como fallas, domos, afloramientos de rocas y semejantes, están asociadas con variaciones apreciables de la morfología del terreno, en concreto de su topografía, rugosidad y capa de vegetación. Los procesos erosivos producen también rasgos y patrones reconocibles. Estos pueden observarse con claridad y, en algunos casos, hasta enriquecerse por la imagen de radar, por la sencilla razón de que el reflejo de las microondas es muy sensible a los cambios de las propiedades físicas de la superficie.

En cualquier longituda de onda, la observación de detalles topográficos depende estrechamente de la geometría de la iluminación. Los detalles se aprecian mejor cuando la iluminación es casi perpendicular a la dirección del curso topográfico. En el caso de sensores del visible y del infrarrojo cercano, la dirección de iluminación viene determinada por la posición del Sol y varía principalmente en la dirección este-oeste. En el caso de los sensores de radar, podemos elegir y controlar la dirección de iluminación y el ángulo de incidencia, lo que confiere mayor flexibilidad. Los cambios de pendiente de unos pocos grados pueden variar el reflejo de radar en un factor de dos o más, con lo que los sensores de radar resultan ideales para detectar los rasgos estructurales pequeños que modifican la topografía. Esta ventaja de las imágenes de radar reviste particular interés en el estudio de regiones de gran densidad forestal, como las de los Trópicos. Para detectar rasgos topográficos menores se han ideado técnicas especiales de procesado de imágenes, mejorando las variaciones espaciales, tanto de baja frecuencia (o de gran escala) como de alta frecuencia (o de pequeña escala).

Los procesos de erosión se producen a muchas escalas. La rugosidad de la superficie, en la escala de unos pocos centímetros a unos cuantos metros, depende de su composición y del proceso erosivo particular a que esté sometida. La rugosidad, en este rango de escalas, tiene un efecto limitado sobre la reflectividad óptica del terreno; es, en cambio, el factor dominante que determina el reflejo del radar. Este efecto ha demostrado su especial utilidad en la cartografía de regiones muy áridas, donde el "barniz del desierto" y las delgadas capas de arena tienden a atenuar la reflectividad y a reducir las variaciones espectrales en las regiones visible y del infrarrojo cercano.

Habida cuenta de la larga longitud de onda de la radiación utilizada por los

sensores de radar, podría esperarse cierta penetración por debajo de la superficie. La profundidad de la penetración suele ser proporcional a la longitud de onda de la señal y guarda una estrecha relación de dependencia con la humedad de la superficie. En las regiones muy áridas del mundo se puede penetrar hasta unos cuantos metros superficie adentro. Hace poco, al analizar los datos recogidos por el *SIR-A* sobre el Sahara Oriental de Egipto, los miembros de la Inspección Geológica de los Estados Unidos y del Jet Propulsion Laboratory se dieron cuenta de que muchas de las características de gran escala que se veían en las imágenes de radar no se distinguían en las del Landsat o desde el suelo. Estas características parecen ser las del subsuelo, cubiertas por una capa de arena de unos metros de espesor. La presencia de esos rasgos ha sido largamente sospechada por los geólogos dedicados a in-





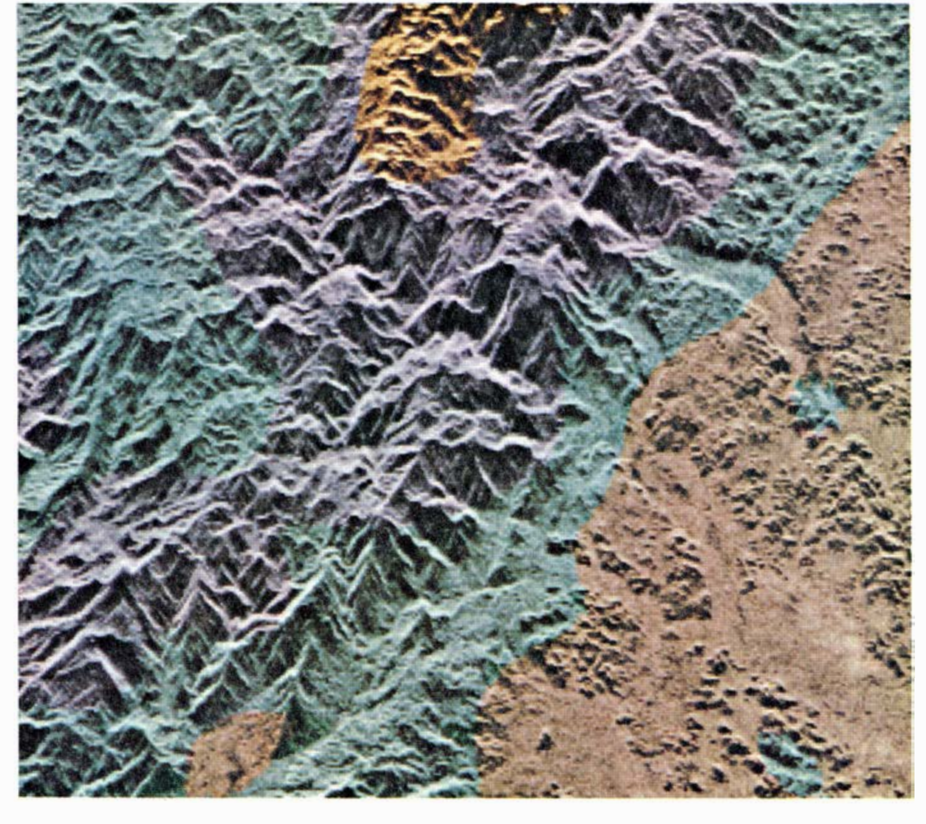
5. VIAS DE DRENAJE de un desierto rocoso de la frontera entre Irak y Arabia Saudí. Apenas visibles en la imagen del Landsat (*arriba*), quedan nítidamente definidas en la imagen del radar *SIR-A* (*abajo*), de la misma zona. Las corrientes de agua intermitentes cortan los valles angostos, llamados wadis, en la roca descubierta. Los wadis están parcialmente rellenos de arena arrastrada por el viento, que refleja una señal mucho más débil al radar que las rocas circundantes. El fuerte contraste que ofrece la imagen de radar facilita el dibujo del patrón de drenaje mejor que con la imagen del Landsat.

vestigar la historia geológica del Sahara. Entre ellas se incluyen valles (algunos de la anchura del actual valle del Nilo), terrazas, cauces desérticos y lechos de roca firme. La capacidad de penetración del suelo de los sensores de radar, recientemente comprobada por investigaciones del suelo, promete tener implicaciones trascendentales en la exploración arqueológica, geológica e hidrológica de las regiones áridas.

También en los húmedos Trópicos se puede traspasar la cubierta vegetal hasta un cierto grado por medio de las señales de microondas. Las imágenes del *SIR-A* del sur de Borneo parecen delinear áreas pantanosas completamente tapadas por la vegetación. A resultados semejantes llegaron miembros de la Universidad de Arkansas al analizar imágenes del *Seasat* de la región central de Arkansas.

Con los sensores remotos se pueden diferenciar e identificar mejor los tipos de rocas si analizamos la "firma" espectral de la radiación visible e infrarroja reflejada o emitida. Los avances recientes deberían extender estas imágenes espectroscópicas bidimensionales a gran número de bandas del espectro visible y del infrarrojo para que sirvan de diagnóstico de la composición de la roca superficial. En esta materia, las imágenes de radar pueden aportar una información complementaria e indirecta a la cartografía litológica. La rugosidad de la superficie y la textura topográfica de pequeña escala, que intervienen de una manera destacada en el eco del radar, dependen del tipo de roca presente. Así, la información de radar es valiosa para diferenciar, aunque no para identificar, los tipos de roca. Por ejemplo, una vez identificada una zona litológica por otros medios, la cartografía radárica ayuda a definir sus límites.

El trabajo realizado por mi grupo en el Jet Propulsion Laboratory ha puesto de relieve que la unión de los datos del radar a los suministrados por las antenas exploradoras ("scanner") multiespectrales del Landsat mejora la capacidad de clasificación de las rocas. Los datos de radar adquieren una importancia incluso mayor en las zonas de vegetación, donde no se pueden hacer imágenes multiespectrales de la superficie. En estas áreas, la única aproximación viable para una somera clasificación de las rocas consiste en aprovechar las variaciones que ofrecen la textura de la superficie y la densidad de drenaje, dos propiedades que suelen estar relacionadas con los tipos de roca. En el Jet Propulsion Laboratory y en la Universidad de Arkansas se están probando algunas técnicas de procesado de imágenes para clasificar zonas, basadas en la textura observada en las imágenes de radar. Estos esfuerzos han superado con éxito la prueba del estudio de regiones de vegetación densa.



6. VARIACIONES DE TEXTURA analizadas matemáticamente para asignar falsos colores a una imagen del *SIR-A* de una región montañosa de Belize. Las diferencias en la textura de distintas partes de la imagen corresponden a variaciones espaciales de las formas superficiales. Thomas G. Farr, del equipo del autor, está investigando esta técnica que permitirá cartografiar las áreas en razón de los tipos de rocas.

La presencia de vegetación y la de construcciones levantadas por el hombre tienen un efecto apreciable en el reflejo radárico. La vegetación tiende a dispersar las ondas de radar, debido, en gran parte, a su alto contenido de humedad y a la densa aglomeración de superficies reflectantes de la cúpula forestal. Los campos de cultivo y los bosques despejados aparecen delimitados con nitidez en las imágenes de radar. En la Universidad de Kansas se trabaja ahora en la solución del problema siguiente: relacionar la intensidad del reflejo de radar con la naturaleza de la vegetación, su etapa de crecimiento y su estado de salud. Cuantos intentos de correlacionar las medidas en el suelo con las imágenes de radar se han realizado hasta el momento, han dado resultados alentadores.

Las obras públicas reflejan intensamente las señales de radar; lo que se explica, principalmente, porque las superficies planas perpendiculares tienden a formar rincones reflectores que devuelven la mayor parte de la energía incidente y porque las estructuras metálicas actúan de antenas que reirradian fuertemente dicha energía. En las imágenes del *Seasat* y del *SIR-A* se recogen muchas construcciones humanas, aun cuando su tamaño no alcance un elemento de resolución. Entre las que se observan con mayor frecuencia están las vías de ferrocarril, carreteras nacionales, torres de alta tensión, puentes, plataformas petrolíferas de sondeo y buques transoceánicos. Se está comprobando la fiabilidad de la combinación de imágenes de radar y del Landsat al objeto de mejorar la capacidad de control remoto de los patrones de utilización del suelo urbano a gran escala.

Todos los sensores de imágenes, los de radar incluido, conducen a representaciones bidimensionales de la imagen. Por tanto, la topografía de la superficie no puede deducirse por entero de una sola imagen. Se necesita un par estereoscópico de imágenes para determinar la altura de los distintos pixeles y generar un plano acotado. El proceso de cartografía topográfica con pares estereoscópicos podría realizarse con las imágenes de radar; y en esa línea de investigación se está. Se trata, empero, de un método tedioso, sobre todo para cartografía de gran escala.

El hecho de que el sensor de radar tenga su propia iluminación controlada permitió a L. C. Graham, de la Goodyear Aerospace, introducir una técnica de interferómetro en conjunción con la de abertura sintética para reunir así información tridimensional de la superficie. Su idea consiste en operar con dos antenas separadas, ambas montadas sobre la misma nave. Sólo una antena ilumina la superficie, pero las dos recogen la señal de retorno. Los dos receptores trabajan con una fuente común de referencia para comparar las fases de los dos ecos recibidos. La información de fase permite medir directamente la altura de cada pixel en la imagen resultante y elimina la necesidad del análisis estereoscópico. Los trabajos más recientes y una simulación detallada de esta idea, realizada por los miembros del Jet Propulsion Laboratory, indican que es factible desde el espacio.

Las señales de microondas no penetran a una profundidad significativa en el agua. Por otra parte, la intensidad del reflejo de radar es muy sensible a las ondas capilares de la superficie y a las pequeñas olas, al igual que lo es a la rugosidad de la superficie terrestre. En virtud de ello, podría detectarse en principio por radar de satélite cualquier rasgo del océano que altere su rugosidad. A su vez, muchas características del océano influyen sobre la rugosidad de su superficie. El oleaje genera una modulación periódica de la intensidad de las olas más pequeñas, pudiéndose así observar normalmente en las imágenes de radar la secuencia del oleaje. La velocidad de las corrientes superficiales también modula la rugosidad, siendo a su vez modulada por la topografía del fondo en las zonas poco profundas, resultando así que se han observado características del fondo de la costa en muchas imágenes de radar, a pesar de que ninguna microonda llega al fondo.



7. IMAGEN DE RADAR SIMULADA del monte Shasta, California, ideada por Michael Kobrick, en el curso de la investigación sobre la fiabilidad de la obtención de información tridimensional de la superficie de la Tierra por medio de un interferómetro de radar espaciotransportado. En este sistema, se mediría punto por punto la diferencia de fase entre los ecos de microondas recibidos por dos antenas montadas sobre la misma nave para generar una imagen topográfica. Aquí, los datos topográficos se obtuvieron realmente, por medios convencionales primero y procesados después, para simular una salida digital de un interferómetro de radar orbital. Los contornos topográficos multicolores, derivados de la corriente resultante de información digital, corresponden a líneas de altitud constante sobre el nivel del mar.

Las ondas internas se propagan a lo largo de estratos de distinta homogeneidad en el océano. La circulación de agua asociada con estas ondas modula la rugosidad de la superficie según patrones espaciales identificables y bien definidos. Las imágenes hechas con el *Seasat* y el *SIR-A* mostraron que las ondas internas eran mucho más comunes de lo que se había pensado. Los torbellinos, anillos de agua caliente, frentes, sedimentos suspendidos y bolsas de aire casi superficiales también afectan a la rugosidad de la superficie y son observados en las imágenes de radar. Estos fenómenos pueden ser controlados prescindiendo de las condiciones climáticas, capas de nubes u hora del día o de la noche. Así, el radar de satélite proporciona una cobertura repetitiva para el estudio de los cambios que se produzcan en la superficie o para rastrear sus movimientos.

Además de los efectos de la rugosidad superficial sobre la intensidad del reflejo de radar del océano, el movimiento de la superficie afecta a la frecuencia de la señal reflejada debido al efecto Doppler. Miembros del Jet Propulsion Laboratory y del Environmental Research Institute de Michigan investigan la explotación de este pequeño corrimiento Doppler para cartografiar la velocidad de las corrientes oceánicas.

Los hielos flotantes, cordilleras de hielo y canales de agua polares tienen un reflejo de radar característico que recoge la imagen de radar. Una cobertura repetitiva del mar de Beaufort con el radar de abertura sintética del *Seasat* en el verano de 1978 permitió que Benjamin Holt, del Jet Propulsion Laboratory, siguiera determinado aspecto de los hielos a más de 160 kilómetros de altitud. Esta investigación posibilitará conocer el movimiento a largo plazo del hielo polar, que está dirigido por las corrientes polares, y que se cree que da una vuelta completa al polo en dos años. Es más, el control continuo de los hielos flotantes es necesario para la planificación de rutas marítimas y la colocación de las plataformas de perforación en los mares polares. Esta capacidad ha alentado al gobierno canadiense a planear un satélite de observación de los hielos con sistema de radar de abertura sintética, que se habrá de lanzar a últimos de los ochenta. La delimitación de la capa de hielo y su extensión es también de particular interés en la determinación del flujo de calor entre el océano y la atmósfera, importante variable de los modelos de clima mundial.

Hasta hace muy poco, los análisis de las imágenes de radar eran principalmente de orden cualitativo. Se basaban en la delimitación de características y figuras y medían su tamaño y estudiaban las asociaciones de formas para inferir alguna información sobre la super-

ficie. En los últimos años se han ido incorporando técnicas cuantitativas que aportarán una mejor comprensión de la imagen radárica de las características de la superficie. Las imágenes de radar calibradas podrían suministrar información adicional acerca de la rugosidad de la superficie y de la humedad del suelo; las técnicas matemáticas se aplicarían a la clasificación de texturas. Los datos cuantitativos, tanto del radar como de las imágenes multiespectrales, podrían ayudar a la identificación de características superficiales. Una serie de imágenes de radar hechas con haces que difieran en sus planos de polarización permitiría separar las medidas de la rugosidad superficial y de la constante dieléctrica (propiedad eléctrica fundamental). Las imágenes hechas con frecuencias múltiples podrían emplearse también para medir la escala de la rugosidad superficial. Muchas técnicas creadas en un comienzo para la interpretación de las imágenes del Landsat se aplican ahora a las imágenes de radar, facilitando el rápido progreso en el desarrollo de los métodos de radar.

Los programas van dirigidos a la creación de sistemas de observación por radar más refinados. En los Estados Unidos se hace hincapié en sensores de radar para investigación múltiple, puestos en órbita por la lanzadera espacial. El próximo paso del programa norteamericano será el vuelo del *SIR-B*, en 1984, que generará imágenes con una amplia variedad de geometrías de iluminación. En Europa, Canadá y Japón, el énfasis se inclina por el desarrollo de sistemas de satélites para observación a largo plazo.

El sensor de radar constituye el único instrumento de que se dispone para observar la superficie de otros cuerpos del sistema solar. Poco se conoce de la superficie de Venus y de Titán (la mayor luna de Saturno) debido a su capa global de nubes. Durante la reciente misión orbital a Venus del Pioneer, un radar cartográfico, con una resolución de unos 80 kilómetros, abrió la primera ventana que nos dejó ver muchas de las principales características superficiales de este planeta. En fase de proyecto, pendiente todavía de aprobación, hay una misión cartográfica de radar a Venus que quizá nos traiga algún día una imagen mucho más detallada de su superficie. Nada sabemos de la superficie de Titán, lo que da pie a la lucubración. No queda otra opción que la del radar de imágenes orbital si queremos obtener una imagen global de la superficie de ese mundo remoto.



8. SE DISTINGUEN ONDAS OCEANICAS, de unos 200 metros de longitud de onda, en la imagen de radar del *Seasat* de las aguas que rodean la isla Kayak, Alaska meridional. Las ondas se refractan y difractan al interactuar con la costa de la isla y con el fondo poco profundo de la bahía Controller.

Desarrollo de mapas y bandas en el cerebro

En el cerebro humano, las células nerviosas dibujan sus relaciones con el mundo exterior en una suerte de mapas que, a su vez, se dividen en bandas. Se investiga la formación de éstas creando ranas de tres ojos

Martha Constantine-Paton y Margaret I. Law

El cerebro de los vertebrados constituye la estructura más compleja de cuantas han desarrollado los organismos. El comportamiento en el microscopio de las neuronas, o células nerviosas, da fe de su versatilidad o capacidad analítica: cada una difiere de su vecina en lo relativo a la forma y a las conexiones que establece, por medio de sinapsis, en el extremo de su fibra nerviosa o axón. Mas, por encima de esa diversidad, queda patente una notable similitud. A medida que se van desvelando los patrones de conexión entre las distintas partes del cerebro, se descubren principios de organización que abonan la esperanza de que los tipos de interacciones neuronales subyacentes a esa complejidad no sobrepasen un número manejable.

Uno de estos principios de organización lo tenemos en la formación de mapas. Los axones que parten de neuronas situadas en una región determinada del cerebro y se proyectan hasta otras células nerviosas de una región diferente suelen reproducir relaciones de vecindad. Resulta así que, si dos neuronas se hallan inmediatas en la primera región, formarán sinapsis en la misma célula, o en células vecinas, de la segunda región: la población diana. Semejante regularidad de la proyección axonal se detectó ya en el siglo pasado. Ahora se ha confirmado ese comportamiento en todas las proyecciones a través de las cuales las señales sensoriales alcanzan el córtex cerebral, en las proyecciones por cuyo medio un área del córtex establece conexión con otra y en las proyecciones por las que las zonas del cerebro involucradas en el control del movimiento actúan sobre los músculos del cuerpo.

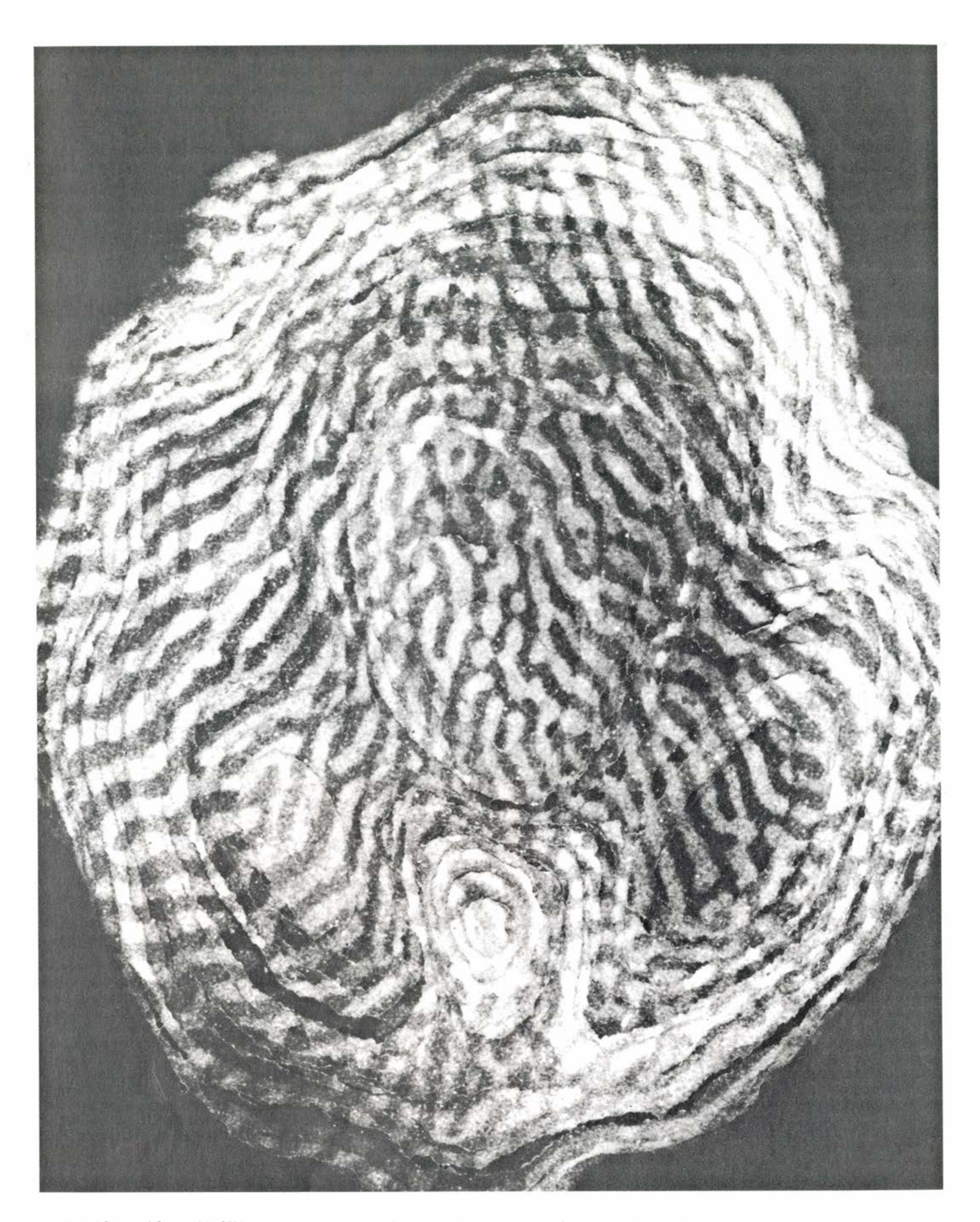
Un segundo principio de organización, descubierto mucho más recientemente, afirma que la división de las regiones del cerebro dibuja un mapa compartimentado en subdistritos periódicos. En este sentido, los trabajos realizados en el labortorio de Jon H. Kaas, de la Universidad de Vanderbilt, han puesto de manifiesto la existencia de una subdivisión, muy regular, del córtex somatosensorial (porción del córtex que recibe aferencias sensoriales de los músculos, articulaciones y la piel). En el córtex somatosensorial, la superficie de una mano aparece representada por un mapa; si tocamos dos puntos cercanos de la piel, se desencadena una actividad eléctrica mensurable en grupos de neuronas que se hallan vecinas en el córtex. En experimentos con monos, Kaas y sus colaboradores han encontrado que, en la capa de la corteza donde los axones aferentes forman sinapsis, la llamada capa 4, el mapa se compartimenta en bandas. Estas separan aferencias sensoriales de la mano de acuerdo con los tipos de información que aporte la aferencia en cuestión. En algunas bandas, las neuronas responden sólo al comienzo de un estímulo; en las bandas afectadas las neuronas ofrecen una respuesta más prolongada. Es como si el mapa de la mano que se dibuja en el córtex somatosensorial del mono se hubiera construido con bandas extraídas, alternativamente, de dos mapas de la mano distintos. Un mapa (y un grupo de bandas) representan los terminales nerviosos cutáneos de adaptación rápida; el otro representa terminales que se adaptan de una manera más lenta.

Edward G. Jones y sus colegas, de la Facultad de Medicina de la Universidad de Washington, demostraron anatómicamente que estas subdivisiones funcionales debíanse a que cada subdivisión recibía axones específicos. Para ello emplearon una técnica radiográfica. Inyectaron en el córtex somatosensorial de un lado del cerebro de mono una pequeña cantidad de aminoácidos marcados con tritio, isótopo radiactivo del hidrógeno. Las neuronas del córtex captaban los aminoácidos y los transportaban, a lo largo de sus axones, hasta el córtex somatosensorial del otro lado del cerebro. Una vez allí, los axones marcados generaban un patrón de radiactividad detectable en los cortes de tejido cubiertos con una emulsión fotográfica sensible a la radiactividad.

El patrón de radiactividad en cortes sucesivos indicaba que los axones terminaban en una serie nítidamente definida. Así, las aferencias que cruzaban de un lado del cerebro al otro, igual que las aferencias que portaban información de la piel, terminaban en bandas en el córtex somatosensorial. La distribución de aferencias al córtex somatosensorial no constituye, sin embargo, el único ejemplo de patrones distribuidos en bandas. Hay otros muchos. Se han encontrado bandas en todas las vías sensoriales, en muchas regiones del córtex y en zonas del cerebro muy diversas: colículo superior, cerebelo y bulbo raquídeo.

Más enigmáticas son unas zonas de sinapsis periódicas que forman bandas funcionales en una región del cerebro que, al propio tiempo, representa un mapa. ¿Por qué están ahí? ¿Por qué establece el cerebro una elaborada separación de diversas aferencias cuando, en última instancia, las aferencias convergerán para formar una representación unificada? Quienes firmamos el artículo hemos realizado una serie de experimentos que dan pie a una respuesta verosímil.

Nuestro trabajo se circunscribe al sistema visual, un grupo de proyecciones que lleva la información visual desde la retina hasta estaciones más centrales del cerebro. Se trata de unas vías dete-



1. BANDAS DEL CORTEX VISUAL de un cerebro de mamífero, la porción de córtex cerebral que recibe información procedente de los ojos. Se muestra aquí aproximadamente la cuarta parte del córtex visual de un hemisferio cerebral de un macaco. Se inyectó en uno de los ojos del animal una pequeña cantidad de aminoácido (prolina) marcado con tritio, isótopo radiactivo del hidrógeno; durante dos semanas, la radiactividad fue progresando por los axones, fibras nerviosas que se arrancan del cuerpo celular de las neuronas, desde el ojo hasta el núcleo geniculado lateral del cerebro, y de allí al córtex visual. Al recubrir los cortes de tejido del córtex visual con una emulsión fotográfica, la radiactividad impresionó la emulsión con bandas claras, que se alternan con bandas oscuras, correspondientes al ojo no inyectado. Cada banda mide aproximadamente 350 micrometros de ancho. La ilustración ha sido proporcionada por Simon LeVay, de la Facultad de Medicina de Harvard; se trata de un montaje realizado con secciones progresivas del tejido.

nidamente estudiadas, lo que ha posibilitado que hoy conozcamos la representación del mundo visual en el sistema nervioso central mejor que la representación de otras modalidades sensoriales. Y debemos mucho de lo avanzado en esa labor a David H. Hubel, Torsten N. Wiesel y sus colegas, de la Facultad de Medicina de Harvard, quienes experimentaron con gatos y monos. Sus conclusiones constituyen un análisis detallado de la relación entre la organización topográfica y funcional. En el gato y el mono, las vías ópticas llevan información hasta la región del córtex cerebral denominada córtex visual. El mapa aquí es binocular: nace de axones que transmiten información de cada uno de los ojos del animal. Hubel y Wiesel han hallado, sin embargo, que al introducir un microelectrodo a través del tejido cerebral de la capa 4 del córtex visual se recoge la actividad eléctrica de neuronas en una secuencia alternante muy regular. Una serie inicial de neuronas responden solamente a destellos de luz en el ojo izquierdo del animal, por ejemplo. Luego aparece una serie de células que responden exclusivamente al ojo derecho y, de nuevo, una serie de células que hacen lo propio con el ojo izquierdo.

Hubel, Wiesel y sus colegas han puesto de relieve, además, que dicha alternancia funcional resulta de la separación de los axones que traen información de cada ojo. Al marcar la vía óptica de cada órgano de la visión con aminoácidos radiactivos, se manifiestan bandas que se distribuyen en un patrón semejante a la piel de una cebra a través de la capa 4 del córtex visual. Cada banda contiene células que responden exclusivamente a un ojo, el izquierdo o el derecho. Las células, de modo alternante, proyectan sus axones a neuronas binoculares de las capas corticales que quedan por encima y por debajo de

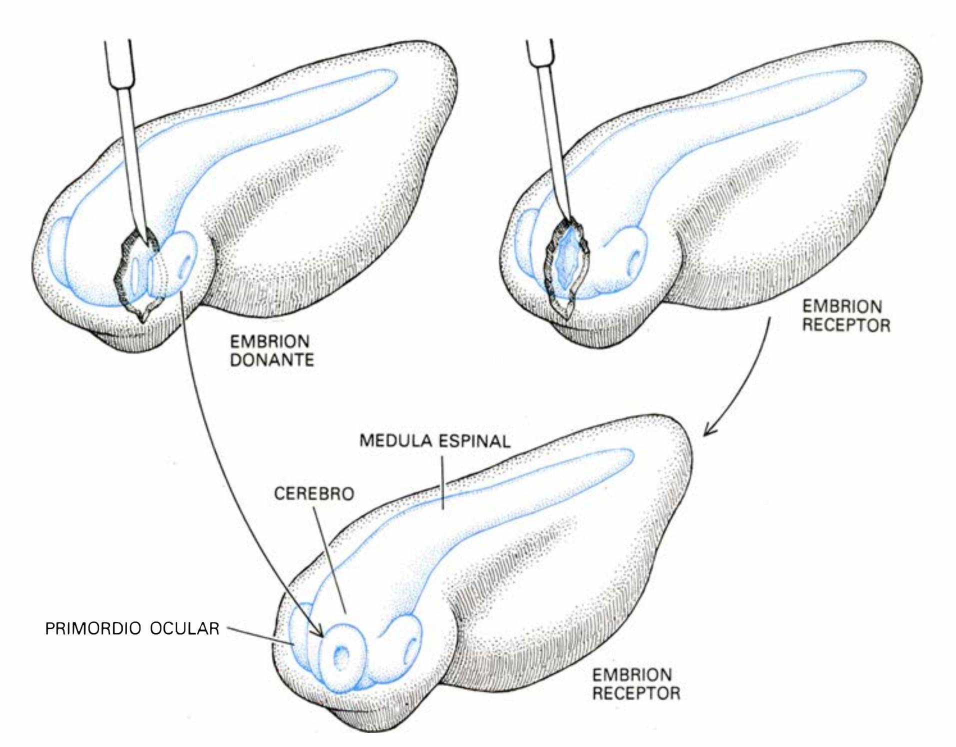




2. BANDAS ANORMALES en el cerebro de una rana leopardo (*Rana pipiens*). Recuerdan llamativamente las bandas del cerebro de un mamífero. Aquí, las bandas se visualizan inyectando el enzima peroxidasa de rábano en uno de los nervios ópticos del animal. El enzima penetra en el cerebro a través de los axones que forman el nervio. El cerebro se trata luego de modo que el enzima produzca, por reacción, una sustancia parda en el interior de los terminales axónicos. La fotografía superior muestra parte del cerebro de una rana normal visto desde arriba. Cada lóbulo constituye un tecto óptico, la región cerebral cuyas células reciben el nervio óptico del ojo situado en el lado opuesto de la cabeza. Se ha inyectado enzima en el nervio óptico del ojo izquierdo, por lo que el tecto de la derecha aparece marcado con el reactivo. La uniformidad del tono pardo sugiere que los axones proyectados al tecto distribuyen sus terminales de forma regular. Las células tectales representan, pues, un mapa topográfico de la retina y, por tanto, un mapa del mundo visual. La fotografía inferior muestra el cerebro de una rana anormal, de tres ojos, derivada del injerto, en el embrión, de un primordio ocular supernumerario (un esbozo de ojo). En esta rana, los axones del ojo supernumerario compiten con los axones de un ojo normal para establecer terminales en el tecto de la izquierda. El enzima se inyectó en los axones supernumerarios, que se aprecian aquí en el tecto, donde sus terminales se distribuyen en regiones en forma de bandas. Las bandas de unos 200 micrometros de ancho, alternan con las correspondientes a los terminales de los axones del ojo normal.

ellas. En otras palabras, cada parte del mundo visual a la que el gato o el mono pueden dirigir sus ojos está cartografiada dos veces en la capa 4: una vez en algún punto de una banda de células que representa al ojo izquierdo y, además, en una banda vecina que representa al ojo derecho.

En nuestros experimentos, llevados a cabo en la Universidad de Princeton, se realizaron ensayos de varias propiedades de una vía óptica bastante diferente, la de la rana leopardo (*Rana pipiens*). Los experimentos se basaron en un procedimiento clásico de trasplante de tejido en embriones de anfibios. El trasplante, sin embargo, se combinó con técnicas modernas de neuroanatomía y neurofisiología para examinar los patrones de las conexiones que establecen las neuronas del sistema visual cuando se colocan en situaciones anormales durante las primeras fases del desarrollo. Ocasionalmente, estos análisis pueden revelar principios de crecimiento y organización no observables a lo largo del desarrollo normal.



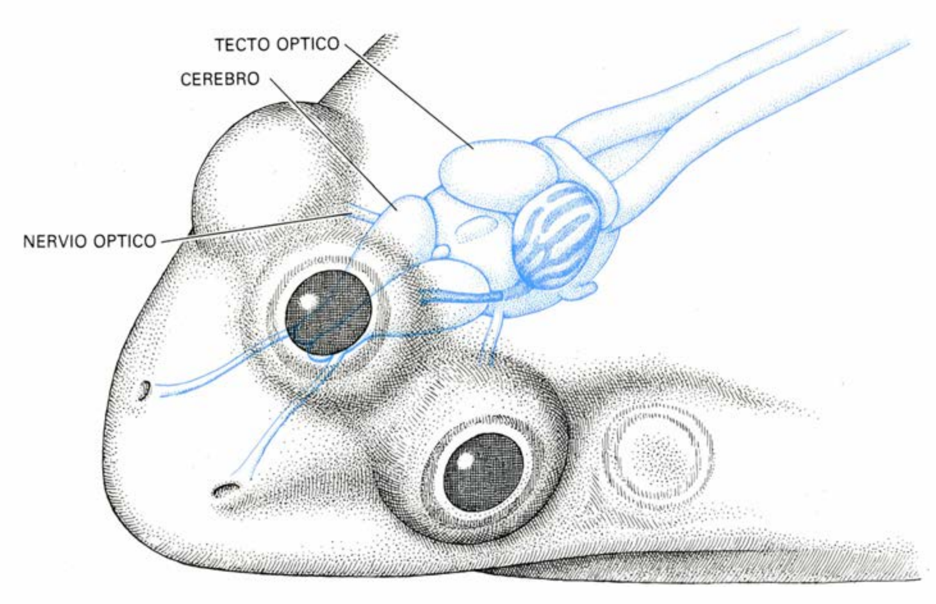
3. PROCEDIMIENTO QUIRURGICO empleado por los autores para obtener una rana de tres ojos. Se toma un primordio ocular de un embrión de rana y se injerta en un segundo embrión tras preparar el tejido para dejar campo al injerto. En el momento del trasplante, los embriones miden unos tres milímetros y los primordios no son más que diminutos botones que sobresalen del cerebro en desarrollo.

En una serie de experimentos extirpamos un primordio ocular (el tejido del que se desarrollará un ojo) de embriones jóvenes cuando el ojo no era más que un botón del sistema nervioso central embrionario. A continuación trasplantamos el primordio a un segundo embrión en la región de sus dos primordios propios. Los embriones tratados de esta forma alcanzaban el estadio de renacuajos y, posteriormente, el de ranas juveniles, con tres ojos bastante normales. El ojo supernumerario solía asentarse frente a uno de los ojos normales. A veces se situaba al final de la nariz, o en la parte superior de la cabeza.

Las ranas leopardos se fían mucho de la visión para su supervivencia, pero, al contrario que los gatos o ratones, sus cerebros no han desarrollado el córtex visual característico de los mamíferos; en ellas, el área más importante para procesar la información visual yace en el tecto óptico, un par de lóbulos simétricos que ocupa gran parte del mesencéfalo. Cada tecto óptico recibe casi todos sus axones retinianos de un solo ojo, el ojo contralateral (el ojo del lado opuesto en la cabeza del animal). La proyección desde este ojo dibuja un mapa muy ordenado de la superficie retiniana; ahora bien, como el tecto no recibe ninguna proyección masiva de la segunda retina, carece de bandas que representen a cada ojo.

Las ranas de tres ojos se apartan de esta descripción. En la mayoría, la retina del ojo supernumerario envía axones predominantemente a uno de los tectos ópticos, donde los axones supernumerarios compiten con las aferencias normales al tecto: los axones que vienen de uno de los ojos normales de la rana. Examinamos el cerebro de las ranas con tres ojos inyectando aminoácidos radiactivamente marcados en el ojo normal o el supernumerario y esperando uno o dos días para que el isótopo viajara a lo largo de los axones hasta sus terminales sinápticos en el tecto. A continuación obtuvimos cortes de los lóbulos tectales de estos animales y los tratamos para así observar qué distribución seguían los terminales sinápticos marcados.

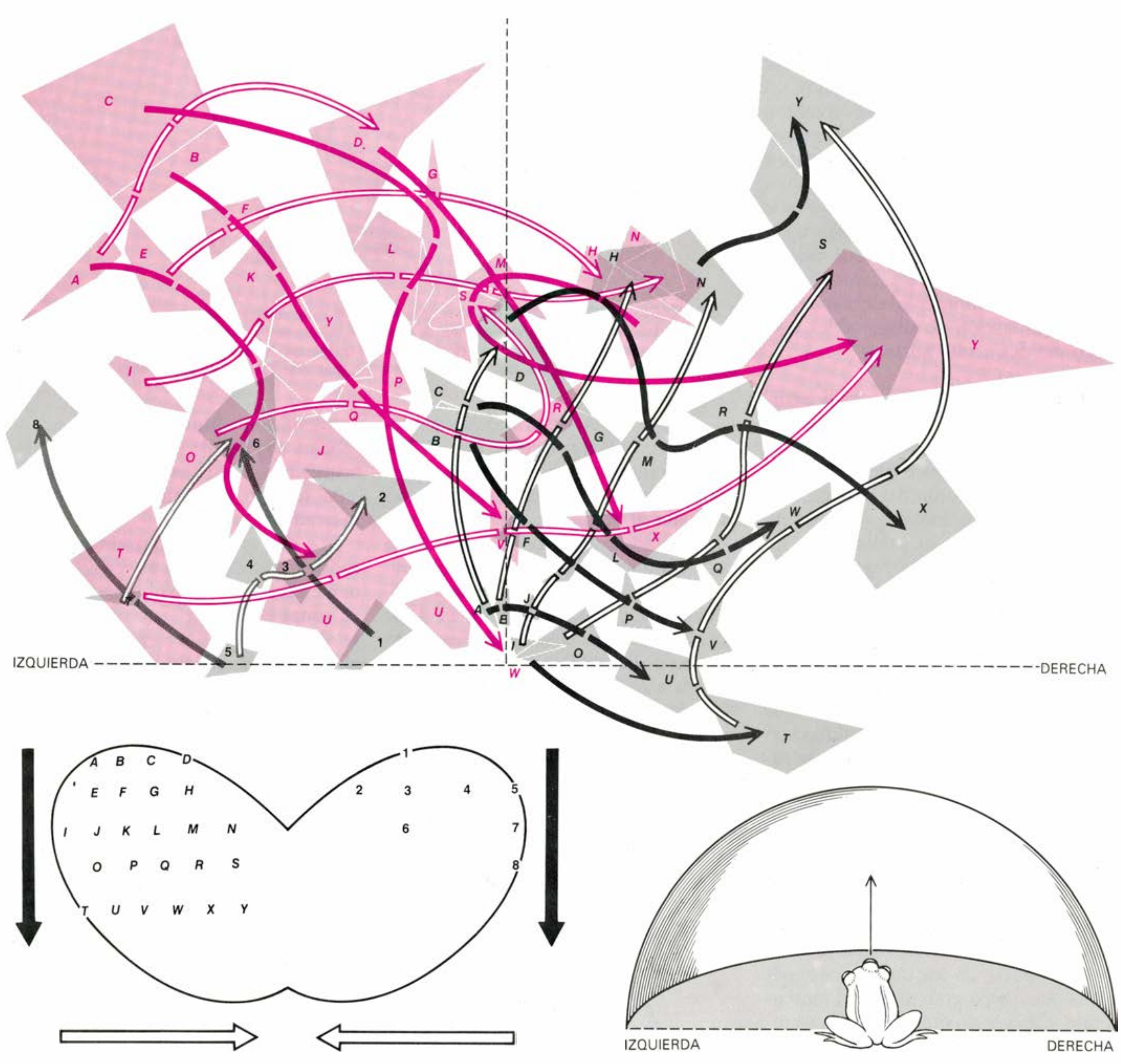
Los dos grupos de terminales (marcados y no marcados) nunca se mezclaban en un corte, sino que se distribuían en zonas específicas para cada ojo, que se alternaban periódicamente. Es más, el rastreo de las zonas de terminales marcadas a través de cortes sucesivos nos demostró que las zonas se alinea-



4. RANA MADURA DE TRES OJOS. Presenta dos ojos normales situados correctamente en la cabeza y un tercer ojo dispuesto frente a un ojo normal o en la parte superior de la testa. En las tres cuartas partes de los casos, el tercer ojo compite con un ojo normal por establecer sus terminales axónicos en un tecto.

ban en bandas. Cada banda medía alrededor de 200 micrometros de ancho y se extendía aproximadamente desde la porción anterior hasta la posterior de cada lóbulo, siguiendo un patrón semejante a la piel de cebra. El patrón era siempre similar, independientemente de que el ojo supernumerario procediera en su origen del lado derecho o izquierdo del embrión donante o de que los axones del ojo supernumerario crecieran en el tecto óptico derecho o izquierdo. La trayectoria tomada por los axones supernumerarios, la forma en que se disponían al crecer y la dirección según la cual entraban en el lóbulo tectal siempre se repetían.

En otro experimento afín, llevado a cabo en varios laboratorios además del nuestro, se extirpó uno de los dos lóbulos tectales de una rana normal o una carpa dorada. Los axones de la retina que se habían proyectado al lóbulo desaparecido regeneraron. Crecieron en el tecto que persistía, donde competían con la proyección preexistente y producían bandas alternantes de terminales que provenían de los axones del ojo izquierdo y del derecho. La duplicación experimental de la aferencia habitual a una región tectal que normalmente contiene sólo un mapa de la retina parece producir inevitablemente, en un



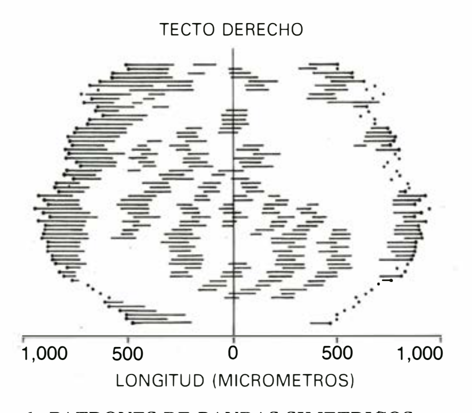
5. ORIENTACION DE LOS MAPAS en los lóbulos tectales de una rana de tres ojos. Se determina registrando la actividad eléctrica de grupos de axones retinianos que terminan en diversos puntos de cada tecto tras disparar un estímulo luminoso desde varios lugares de una superficie hemisférica situada frente a la rana. Los números del *1* al *8* indican los puntos donde se efectuó el registro en el tecto del lado derecho del cerebro. Señalan también los campos de recepción: las porciones del mundo visual que están bajo el control de las situadas en esos campos. En este caso, el tecto del lado derecho del cerebro representa un mapa del mundo que se ve a través del ojo del lado izquierdo de la cabeza (*gris*). Las flechas huecas del mapa enlazan las direcciones del mundo visual, controladas por los terminales axónicos dispuestos en sentido horizontal en el tecto; las flechas rellenas enlazan las direcciones controladas por los terminales axónicos dispuestos de delante atrás. De la *A* a la *Y*, las letras indican los lugares de registro en el tecto del lado izquierdo del cerebro. Señalan también los campos de recepción. En el tecto del lado izquierdo se distinguen dos mapas, el que corresponde al ojo derecho normal (*negro*) y el del ojo supernumerario (*color*). En la mayoría de los lugares de registro, los terminales axónicos que representan a cada ojo se sitúan muy próximos, lo que permite su detección con un único electrodo. Ambos mapas están bien organizados, en el sentido de que al cubrir uno de los dos ojos, la actividad de las células de una hilera de lugares del tecto puede dispararse mandando estímulos desde una hilera de zonas del mundo visual. Sin embargo, el mapa normal y el mapa supernumerario no encajan: cuando el animal abre los dos ojos, los grupos vecinos de células tectales pueden responder a zonas del mundo muy distintas.

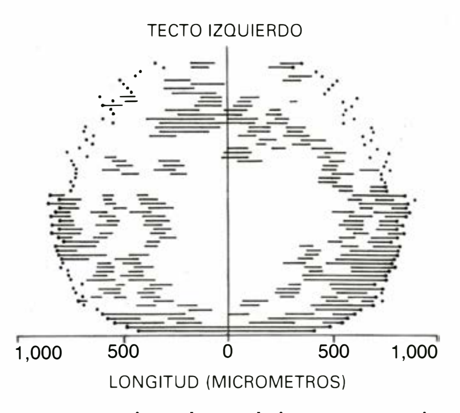
animal vertebrado inferior, un grupo de complejas subdivisiones funcionales marcadamente similares al patrón del córtex visual de los mamíferos normales.

En algunas de nuestras ranas con tres ojos registramos la actividad de los terminales de los axones retinianos en el tecto mientras se lanzaban destellos de luz sobre una pantalla instalada delante del anfibio. Cubrimos primero los ojos normales y, posteriormente, el ojo supernumerario. Demostramos así que cada representación de un ojo sobre el tecto doblemente inervado estaba adecuadamente alineada con respecto al eje original del ojo del embrión. Por tanto, los axones de cada proyección mantenían en el tecto un mapa de la retina de la que procedían, incluso aunque el mapa quedara interrumpido por las bandas interdigitadas. De nuevo las ranas anormales se asemejaban a los mamíferos normales. Dos aferencias retinianas producían dos representaciones distintas del mundo visual en una sola estructura diana.

En los mamíferos, ambos ojos se disponen simétricamente en la cabeza. No así en la rana con tres ojos, donde el ojo supernumerario se ubica de forma anormal en la cabeza. En ésta, la proyección desde la retina supernumeraria hacia el tecto suele transmitir una visión del entorno del animal que no se ajusta a la visión de la retina normal. Ello significa que las neuronas próximas entre sí, y en bandas adyacentes, del tecto reciben información acerca de partes no relacionadas del espacio visual. La situación podría simularse ajustando un ojo del lector a un prisma que desviara la luz hacia ese mismo ojo 90 grados, por ejemplo. Al mirar alrededor, el prisma trasmitiría a un ojo imágenes de lo situado sobre su cabeza, mientras el otro ojo miraría al frente. Ambas imágenes incidirían simultáneamente en la misma parte del córtex visual. El mundo carecería de sentido para nuestro hipotético sujeto.

Si a una rana con tres ojos le presentamos un objeto que va de un sitio para otro sin rumbo fijo, y que simula su presa (un insecto volando), lo normal es que no se inmute. Tal vez dé palos de ciego contra el estímulo. Si le permitimos que mire a través de sus ojos normales, tapándole el supernumerario, los ataques serán certeros. Presumiblemente, los mapas tectales que representan los ojos normales se alinean correctamente con las vías motoras del sistema nervioso que controlan el comportamiento del animal. ¿Qué ocurre si sólo tiene libre el ojo supernumerario?





6. PATRONES DE BANDAS SIMETRICOS en una rana con tres ojos en la que el ojo supernumerario envía axones a ambos tectos ópticos. En concreto, un hueco en el patrón de bandas de un tecto se corresponde con una zona simétrica de bandas en el otro. Los patrones sugieren que los axones del ojo supernumerario compiten en el tecto con los axones de un ojo normal solamente en ciertas zonas del tecto, que vienen determinadas por el lugar donde nacen los axones retinianos. La superficie de cada tecto se reconstruyó midiendo la anchura de las bandas en una serie de cortes realizados a intervalos de 20 micras.

Volverá a fallar el golpe. Las vías motoras conducidas por el mapa visual mal alineado ordenan el cuerpo del animal en direcciones que no se corresponden con la posición de la presa en el espacio.

No tiene sentido preguntarse por la significación funcional o evolutiva de las bandas en el cerebro de las ranas con tres ojos. El tercer ojo es anormal y, en ausencia de aferencia importante de los dos ojos normales a un solo tecto óptico, la rana normal no recibiría ningún beneficio de un mecanismo desarrollado específicamente para separar las aferencias tectales en bandas. Por otro lado, la supervivencia de las ranas en libertad y, en particular, su capacidad para cazar insectos de los que se alimentan, depende críticamente de un mecanismo vigoroso que asegure la formación de un mapa preciso de la superficie retiniana contralateral en cada tecto.

Comenzamos, así, a considerar la posibilidad de que las bandas aparecieran por el mismo mecanismo de desarrollo que genera los mapas. Simon LeVay, que a la sazón colaboraba con Hubel y Wiesel en la Facultad de Medicina de Harvard, sugirió por primera vez, en 1975, esa relación. LeVay propuso que las bandas funcionales del córtex visual del ratón representarían un compromiso entre dos tendencias opuestas: un proceso de difusión, en el que los axones portadores de información de cada una de las retinas tratarían de llenar todo el córtex visual con un mapa, y un proceso de agrupamiento, en el que los axones portadores de información de cada una de las retinas tratarían de permanecer unidos, como si fueran repelidos por las aferencias del otro ojo. El resultado más probable de las dos tendencias opuestas serían las bandas interdigitadas, porque una configuración en bandas optimiza simultáneamente ambos procesos.

¿Cuáles son, entonces, los mecanismos que dan origen a los mapas neurales? ¿De qué forma generan bandas estos mecanismos cuando dos poblaciones de axones se representan en una única zona diana? Afortunadamente, las proyecciones desde la retina hasta el tecto óptico de los vertebrados inferiores han sido objeto, desde hace tiempo, de estudios de cartografía neural. Entre los primeros de este campo se contaba R. W. Sperry, del Instituto de Tecnología de California. Sperry rotó quirúrgicamente los ojos de tritones 180 grados. En algunos de los animales dejó intactos los nervios ópticos; en otros, los extirpó y los dejó regenerar. En uno u otro caso los tritones cometían errores de 180 grados cuando querían alcanzar el estímulo, errores que no se corregían con el paso del tiempo. Los animales se comportaban como si ignoraran que se había rotado sus retinas. Evidentemente, cada parte de la retina continuaba proyectando sus axones a una región determinada del tecto, a pesar de que Sperry había alterado la situación de forma que cada parte de la retina recogiera ahora zonas anormales del mundo visual. Numerosos estudios ulteriores ampliaron la labor de Sperry, demostrando que la parte del tecto inervada por una determinada zona de la retina está prefijada en el embrión incluso antes de que los axones abandonen la retina y alcancen el cerebro en desarrollo.

En 1963 Sperry avanzó una hipótesis para explicar el alineamiento de los mapas visuales en el cerebro. Sugirió que las células retinianas y las células tectales se desarrollaban de acuerdo con su

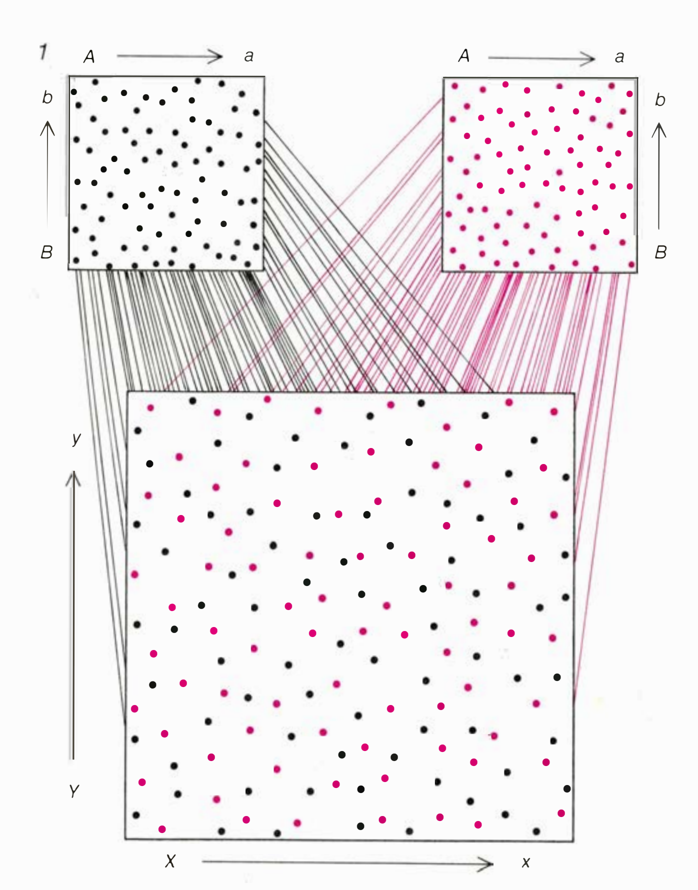
posición a lo largo de dos ejes, en la retina y en el tecto, respectivamente; cada célula venía a tener en su superficie un grupo único de moléculas marcadoras. Los axones de las células retinianas podrían entonces hacer sinapsis solamente con las células tectales que poseyeran los marcadores complementarios. En resumidas cuentas: Sperry imaginó un emparejamiento por afinidad química, bastante rígido, entre retina y tecto. El emparejamiento, sin embargo, no puede ser absoluto. Experimentos realizados en peces y anfibios en varios laboratorios han demostrado que, bajo determinadas condiciones, los axones forman sinapsis con neuronas tectales que no suelen ser sus destinos normales. Una retina reducida a la mitad de su tamaño mediante intervención quirúrgica puede enviar sus axones a formar una proyección extendida a todo un lóbulo tectal. Y a la inversa, los axones de toda una retina pueden comprimir su mapa de suerte que ocupen un tecto reducido a la mitad.

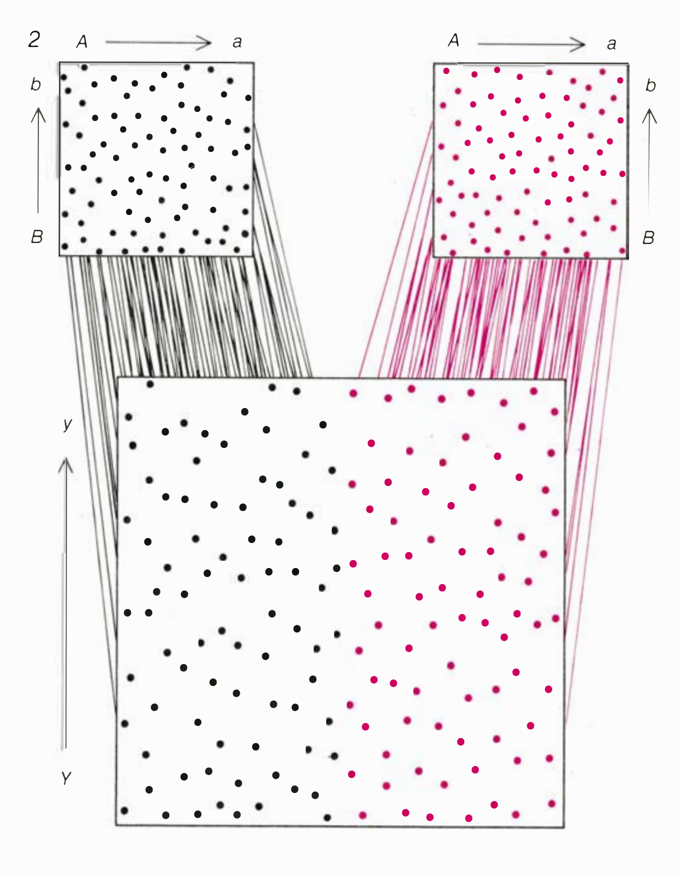
Un mecanismo basado en un emparejamiento rígido de marcadores fijados en las células retinianas y tectales no explica la plasticidad que se aprecia al variar quirúrgicamente los tamaños de las poblaciones retiniana y tectal. Las posiciones de los axones retinianos en el tecto deben estar controladas por algún mecanismo que pueda adaptarse a los cambios en el número relativo de células retinianas y tectales.

¿Cómo se realiza esta adaptación? Se han defendido tres posibilidades. Primero, la rígida quimioafinidad por marcadores propuesta por Sperry sería capaz de "reespecificación", de modo que la intervención quirúrgica induciría un cambio del marcador en la retina o en el tecto. Segundo, en vez de venir condicionada por muchos marcadores diferentes, la identificación de células en la retina o en el tecto dependería del gradiente de dos sustancias marcadoras, una por cada eje.

Según la tercera posibilidad, la retina y el tecto carecen de marcadores. En su lugar, los axones que se proyectan desde la retina al tecto mantendrían su ordenación relativa gracias a su cohesión, que persistiría según crecen hacia el tecto. Se debe explicar entonces por qué el mapa, como un todo, viene siempre a tener la misma orientación en el tecto. En particular, los mapas retinianos de todos los vertebrados no mamíferos representan la parte central del mundo visual del animal en la porción anterior del lóbulo tectal y las partes más laterales del mundo visual en las porciones más posteriores del lóbulo.

Se dispone ya de pruebas suficientes para calibrar el valor de esas posibilidades. La primera posibilidad, la idea del cambio o reespecificación de los rígidos marcadores retinianos o tectales, no se sostiene. Manipulaciones quirúrgicas realizadas en la dorada en varios laboratorios han demostrado que el tecto puede recibir secuencialmente impulsos de una retina normal, luego de una hemirretina expandida y luego, de nuevo, de una retina normal. Por otra parte, informes de experimentos con ranas del género *Xenopus* indican que una región del tecto (aunque posiblemente no las mismas células tectales) recibe, a la vez, aferencias de la mitad embriológicamente anterior de una retina y de la mitad embriológicamente posterior de otra retina. Si las marcas tectales se reespecifican, podrán hacerlo a menudo. Es más, las células de una pequeña





7. COMPARACION DE POSIBLES MECANISMOS de desarrollo de mapas en el cerebro. En la zona superior de cada sección de la ilustración aparecen dos retinas. Sus células (*puntos*) se suponen marcadas por gradientes de concentración de dos sustancias (*A*, *B*) localizadas en la superficie celular. En la parte inferior de cada sección de la ilustración se halla un tecto al que envían sus axones las retinas. En la forma más simple del mecanismo denominado emparejamiento por quimioafinidad (*1*), las células del tecto estarían marcadas por gradientes (*X*, *Y*), cuya complementaridad con los marcadores retinianos guiaría a los axones en crecimiento. Los axones de ambos ojos se mezclarían al establecer terminales en el tecto, situación nunca observada en las ranas de tres ojos. En otro mecanismo algo diferente (*2*), los axones de cada retina mantendrían su ordenación espacial en el crecimiento hacia el

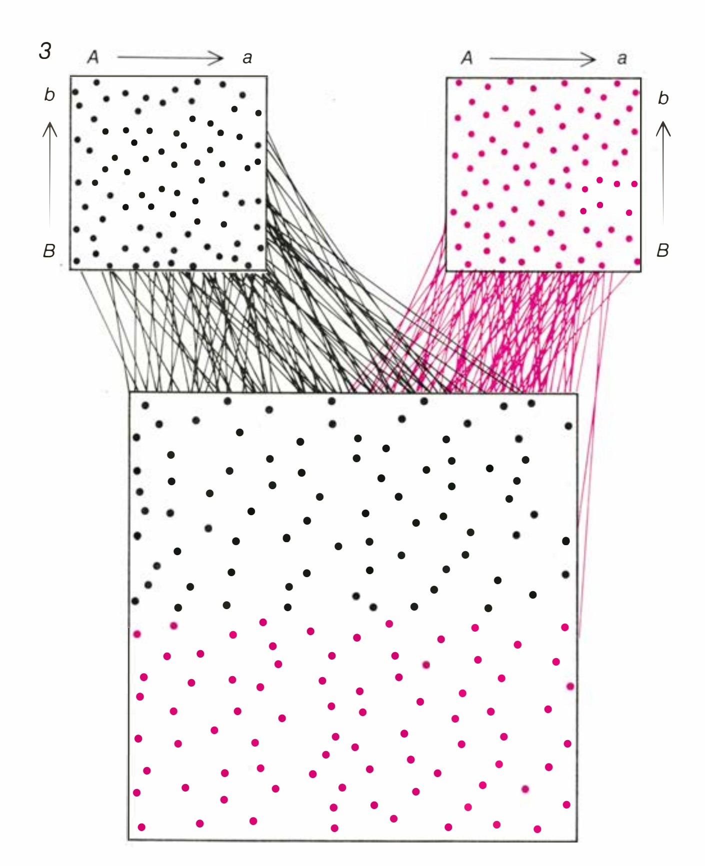
región del tecto cambian sus marcas sin previo acuerdo con sus vecinas. Los marcadores tectales serían tan plásticos que no conseguirían identificar una célula por su posición tectal.

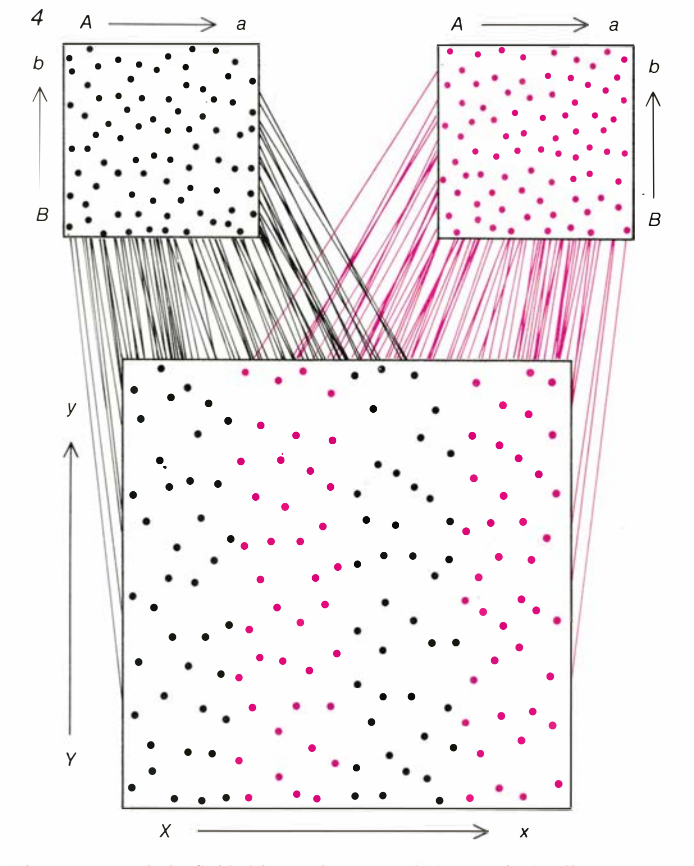
Por su parte, los marcadores de la retina, si existen, no parecen mutables. Scott E. Fraser, trabajando en la Universidad Johns Hopkins, extirpó un ojo de renacuajo de *Xenopus*. El ojo intacto proyectó axones al tecto contralateral, igual que lo hubiera hecho normalmente. Además, la porción inferior del ojo intacto, cuya contribución al nervio óptico estaba todavía en fase de desarrollo en el momento de la intervención quirúrgica, emitió axones hacia la totalidad del tecto ipsilateral, que hubiera sido inervado por el ojo extirpado. Si una región inferior de la retina puede proyectar un mapa expandido a un tecto y un mapa normal al otro, debe conectarse a células en diferentes posiciones tectales. Es poco probable, pues, que en la expansión de un mapa participe la reespecificación de marcadores retinianos.

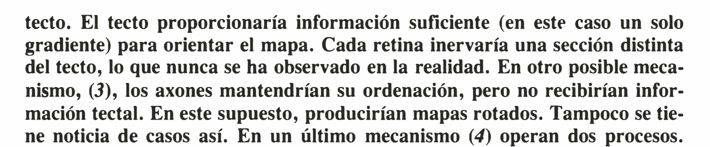
La segunda posibilidad, la de una gradación de marcadores en la retina y el tecto, explica las expansiones o compresiones de un mapa. Justifica también la capacidad de las células de una región dada del tecto para recibir axones de diferentes partes de una retina, así como la facilidad con que una porción de retina puede enviar axones a regiones bastante diferentes de dos lóbulos tectales. Basta un gradiente de un marcador molecular a lo largo de cada uno de dos ejes para proporcionar, a cada posición tectal y retiniana, una combinación única de marcadores; el emparejamiento entre las posiciones tectal y retiniana puede ajustarse a la extensión de los marcadores de la retina o el tecto.

La hipótesis de la gradación de marcadores predice, sin embargo, que la orientación de una proyección retiniana sobre el tecto se mantiene después de cualquier perturbación. Varios experimentos demuestran lo contrario. Ronald L. Meyer, del Instituto de Tecnología de California, extirpó la mitad de la retina de una dorada. Con la intervención se eliminó la aferencia hacia la mitad de un lóbulo tectal. Al mismo tiempo, Meyer forzó a la mitad de los axones del otro ojo a proyectarse hacia el tecto semilibre. Cabía predecir que los axones reconducidos terminarían en el tecto semilibre de forma semejante a como se proyectan normalmente. En este experimento, los axones que Meyer recondujo habrían terminado en la parte del tecto semilibre que retuvo su aferencia retiniana. Sin embargo, los axones reconducidos formaron una proyección alterada (de hecho invertida) en la mitad libre del lóbulo. Aparentemente, los axones reconducidos desde el centro del ojo intacto se aproximaron a su posición tectal apropiada. Los otros axones reconducidos, sin embargo, no podían preservar ni la continuidad ni la orientación de su mapa retiniano, pues ello les hubiera forzado a terminar en la parte ocupada del tecto. Los estudios de Meyer indican que el mantenimiento de las relaciones de vecindad debe suponer una importante tendencia que opere independientemente en la formación de un mapa. Después de todo, en los experimentos de Meyer las relaciones de vecindad se mantenían en una región inapropiada del tecto, a expensas de la orientación normal del mapa.

Estos resultados parecen abonar la






tecto. El tecto proporcionaría información suficiente (en este caso un solo gradiente) para orientar el mapa. Cada retina inervaría una sección distinta del tecto, lo que nunca se ha observado en la realidad. En otro posible mecanismo, (*3*), los axones mantendrían su ordenación, pero no recibirían información tectal. En este supuesto, producirían mapas rotados. Tampoco se tiene noticia de casos así. En un último mecanismo (*4*) operan dos procesos. Primero, una quimioafinidad imprecisa empareja los terminales dispersos en el tecto en una orientación adecuada. (En el sistema visual de la rana leopardo bastaría con que cada axón llegara al cuadrante adecuado.) Luego, un grupo de interacciones locales mantiene la vecindad en el tecto sólo en los terminales de los axones que proceden de células vecinas en una de las retinas. Se generan bandas porque solamente éstas optimizan los dos procesos.

tercera posibilidad, que aboga por una cohesión entre los axones que crecen hacia el tecto y propone que no existe afinidad química entre los axones retinianos y las células tectales. Los defensores de la idea citan estudios que indican que en los peces, ranas y pollos, los axones de muchas zonas de la retina (aunque no todas) crecen hacia el tecto junto con axones de células que son sus vecinas en la retina. Como hemos señalado, sin embargo, la idea no explica por qué la orientación de los mapas normales se mantiene de forma coherente. También es difícil reconciliar la ausencia de marcadores retinianos y tectales con un amplio número de experimentos en los que los axones retinianos rompen la continuidad de su mapa, para terminar, como cabría esperar, en una zona del tejido tectal rotado o trasplantado a una posición anormal en el tecto.

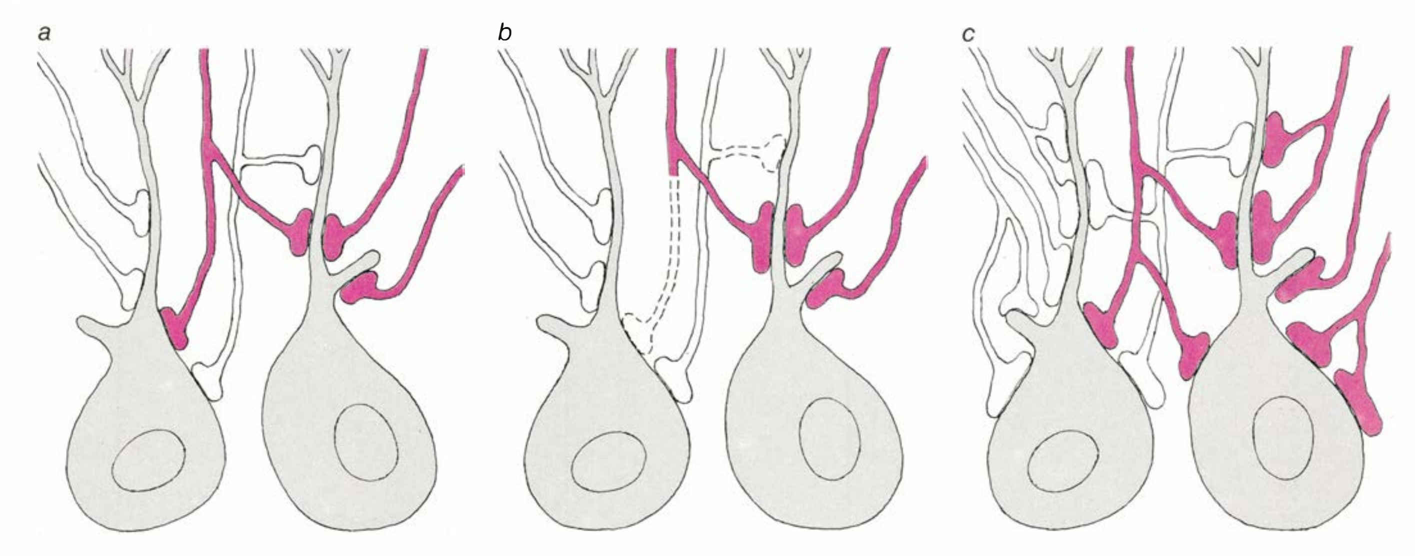
Lo que más nos persuadió de la existencia de marcadores retinianos y tectales fue cierto descubrimiento que observamos en las ranas con tres ojos. En la cuarta parte de las ranas, más o menos, la retina supernumeraria enviaba axones a ambos lados del cerebro; ningún tecto se hallaba, pues, completamente dispuesto en bandas. Vimos que, en estas ranas, un hueco en el patrón de revestimiento de un tecto correspondía a una zona de bandas en la localización simétrica del otro tecto. Si los axones que crecen desde un tercer ojo simplemente conservan la topología de la retina, las proyecciones al tecto debían haberse expandido y formado bandas en toda la extensión de ambos lóbulos. Pero los axones supernumerarios compiten, manifiestamente, con los axones de la retina normal por el espacio tectal sólo en las localizaciones apropiadas para la zona de la retina de la que proceden los axones. Parece, pues, que las células tectales están marcadas y que los axones retinianos distinguen entre las marcas tectales.

Por otro lado, trabajos recientes de nuestro laboratorio revelan que un tecto, al que nunca inervara una retina, podía desarrollar un mapa. Extirpamos los dos primordios oculares de embriones de rana antes de que hubieran comenzado a enviar axones al cerebro. Seguimos luego las proyecciones axonales por las que los tectos habían establecido mapas en otras partes del cerebro. Los mapas eran idénticos en las ranas sin ojos y en las normales. Así pues, las células tectales eran capaces de expresar sus identidades de posición con independencia de sus conexiones con la retina. Obviamente, el tecto debe disponer de alguna forma de información para asegurar el registro más adecuado del mapa retiniano con otros mapas visuales del cerebro.

Se tiene el convencimiento de que ningún mecanismo simple de representación resolverá las controversias que plantean muchas observaciones experimentales. Por un lado parece que los axones de una determinada parte del ojo son capaces de elegir determinada parte del tecto. Por otro, numerosos estudios (y desde luego el centrado en las bandas en el tecto óptico doblemente inervado de las ranas con tres ojos) reflejan una cohesión entre los terminales sinápticos que representan una retina, que escapa a cualquier mecanismo de quimioafinidad entre retina y tecto.

Si suponemos que en el establecimiento de mapas neurales intervienen dos mecanismos independientes, se disipan muchos enfrentamientos. La disposición en bandas aparece, entonces, como una prolongación lógica de los mapas. Admitamos que en las primeras fases del desarrollo de un mapa un gradiente de quimioafinidad dispuesto a lo largo de al menos dos ejes de la retina y del tecto guía el crecimiento de los axones. La conducción no necesita ser de suma precisión: los gradientes podrían resultar suaves y las afinidades bastante débiles. En el sistema visual de la rana leopardo bastaría con que cada axón llegara al cuadrante apropiado del tecto. La precisión del mapa resultaría de un segundo estadio del desarrollo, en el que las interacciones en el tecto mantendrían vecinos a sólo los axones terminales que nacieron de células inmediatas en la retina. El resultado de esta secuencia será el equilibrio descrito por LeVay, Hubel y Wiesel, en el que la zona diana de dos proyecciones queda dividida en bandas terminales alargadas.

El atractivo que rodea a la hipótesis de la quimioafinidad ha instado a los investigadores a buscar moléculas marcadoras en la superficie de las células de la retina y el tecto. Las moléculas deben distribuirse a lo largo de la retina o



8. PROCESOS CELULARES supuestamente subyacentes al desarrollo de un mapa tectal preciso, representados para dos células de un tecto inervado por dos ojos. Un episodio inicial de emparejamiento por quimioafinidad deja imbricaciones locales en las que algunos axones procedentes de los dos ojos inciden sobre las mismas células tectales (*a*). Seguidamente, los terminales que corresponden a uno de los ojos se refuerzan (en el dibujo se representan aumentados) a expensas de los terminales del otro ojo (*b*). Además, la densidad de terminales "correctos" podría aumentar (*c*). En cualquier caso, cada retina acaba dominando ciertos grupos de células. Los terminales que representan células vecinas en una retina probablemente transmiten señales muy correlacionadas. Esta correlación fundamentaría el reforzamiento de las conexiones bien cartografiadas y sería la responsable de la formación de bandas.

el tecto con un gradiente y una capacidad para unirse a otras sustancias que ocasionarían el conocido alineamiento del mapa tectal. Varios avances recientes prometen soluciones. Así, en el laboratorio de Marshall W. Nirenberg, del Instituto Nacional del Corazón, Pulmón y Sangre de los Estados Unidos, exponen células del sistema inmune de ratón a extractos de retina de pollo. Posteriormente, las células del bazo del ratón que fabrican anticuerpos se aíslan y se clasifican. Cada uno de los cultivos de células resultantes fabrica un anticuerpo altamente específico; uno de los anticuerpos obtenidos por esta técnica se une de forma gradual a las células situadas a lo largo de un eje de la retina. Cabe deducir que la desconocida molécula a la que se une el anticuerpo tiene una distribución gradual similar. Desde un enfoque experimental distinto, Willi Halfter, Michael Claviez y Uli Schwartz, del Instituto Max Planck de Investigación Vírica de Tübingen, investigaron la adhesión entre la retina y el tecto. Observaron que los axones de diferentes zonas de la retina del pollo muestran, correspondientemente, diferencias en su capacidad para unirse a membranas aisladas de células tectales de pollo.

Quedan aún por resolver, sin embargo, cuestiones fundamentales del segundo estadio de la formación de mapas: las interacciones que mantienen unidos los terminales de células vecinas y que presumiblemente dan lugar a las bandas. Michael P. Stryker, de la Facultad de Medicina de la Universidad de California en San Francisco, ha demostrado que la tetrodotoxina, fármaco que bloquea la capacidad de las neuronas para comunicarse por medio de potenciales de acción, impide o retrasa el desarrollo de bandas en el córtex visual de los gatos al inyectarla en los ojos de individuos jóvenes. Las proyecciones de los ojos aparecen mezcladas en su zona diana cortical. N. V. Swindale, de la Universidad de Cambridge, ha comunicado resultados similares en gatos criados en la oscuridad.

Aparentemente, la actividad neuronal es esencial para la cohesión entre los terminales sinápticos que representan a uno u otro ojo. ¿Cómo se realiza el proceso? En una retina dada, las células vecinas que proyectan sus axones al tecto (o hacia el córtex visual) tienden a generar secuencias similares de potenciales de acción, porque están conectadas (por medio de neuronas retinianas intermediarias) a muchas de las mismas células receptoras de luz. Además, los potenciales de acción correlacionados de neuronas que son vecinas en la retina tienen más probabilidad que las señales no correlacionadas de inducir actividad eléctrica en una determinada célula tectal. Por tanto, una actividad muy correlacionada en distintos pares de sinapsis destacaría en el tecto (o en el córtex visual) las sinapsis de aquellas células que fueran vecinas en la retina. Si las neuronas tectales reforzaran las sinapsis de las distintas neuronas muy correlacionadas a expensas de sinapsis cuya actividad fuera relativamente ineficaz, el mapa topográfico, en principio burdo, iría ganando precisión.

En síntesis, el modelo con dos mecanismos de desarrollo del mapa tectal propone la existencia de afinidades de débil gradación que alinean aproximadamente a los axones retinianos en el tecto. El mapa se ordena posteriormente de forma precisa por el reforzamiento de sinapsis de células retinianas vecinas que tienden a activarse simultáneamente. Cristoph von der Malsberg y David Willshaw, del Instituto Max Planck de Química Biofísica de Göttingen, han efectuado simulaciones con ordenador en las que el reforzamiento selectivo de sinapsis actúa sobre dos proyecciones topográficas burdas de una única zona diana. En la simulación se forman mapas con bandas.

La idea de que la eficacia de los terminales sinápticos puede determinar su estabilidad y su posición en el cerebro no es reciente, ni se limita a los mapas. En la década de 1940, D. O. Hebb, de la Universidad McGill, sugirió que el reforzamiento selectivo de las sinapsis quizá sea la base de algunos aspectos de aprendizaje. Algunas variantes de la hipótesis de Hebb pretenden explicar el desarrollo de circuitos neuronales en el cerebelo, la especificidad de las neuronas sensitivas a determinados estímulos y la maduración de las conexiones motoras entre el sistema nervioso y los músculos. Aunque los neurólogos están aún lejos de descubrir los mecanismos moleculares que subyacen al reforzamiento selectivo o la estabilización de las sinapsis, el concepto en sí mismo es útil para comprender cómo puede influir la actividad neuronal en la estructura neural. Los potenciales de acción y la eficacia relativa de las señales sinápticas probablemente constituyan el nexo de unión entre los mapas y las bandas del cerebro. Podrían, en efecto, poner en sintonía un sistema nervioso en desarrollo.

Ciencia y sociedad

Genes y desarrollo

Es bien conocido que el desarrollo de los organismos está dirigido por el programa genético contenido en el ADN de sus células. Sin embargo, sólo durante los dos o tres últimos años se ha podido comenzar a estudiar este programa a nivel molecular, al haberse aislado varios genes implicados en el desarrollo. Esto se ha llevado a cabo en la mosca del vinagre, *Drosophila melanogaster*, mediante una feliz conjunción de conocimientos genéticos clásicos y técnicas modernas de recombinación del ADN *in vitro*.

Así, el análisis genético de *Drosophila* ha puesto de manifiesto la existencia de numerosos genes implicados en el desarrollo, como, por ejemplo, los genes del complejo *bithorax*. *Drosophila*, como todos los insectos y otros muchos *phyla* del reino animal, tiene el cuerpo dividido en segmentos. Los genes *bithorax* controlan la diferenciación de los segmentos segundo y tercero del tórax y la de los segmentos abdominales. La función de estos genes la pone de manifiesto el hecho de que las mutaciones que ocurren en los mismos transforman uno o varios segmentos en otros generalmente situados más adelante en el cuerpo del insecto. De esta manera, las moscas mutantes *bithorax-postbithorax* poseen cuatro alas en lugar de las dos características, porque en ellas el tercer segmento torácico se desarrolla de forma parecida al segundo, el único que lleva alas en las moscas normales. El análisis genético también ha revelado que es rasgo común, aunque no exclusivo, de los genes implicados en el desarrollo hallarse organizados en complejos génicos subdivisibles en varios componentes distintos (pseudoalelos), aunque estrechamente relacionados en su función. Además, suelen ocupar regiones bastante extensas del genoma, en posiciones determinadas con gran precisión dentro de los cromosomas politénicos gigantes de las glándulas salivares larvarias. Semejante determinación se ha hecho posible gracias a que dichos cromosomas poseen una morfología típica resoluble en bandas (unas 5000) perfectamente identificables al microscopio óptico.

En organismos superiores, para aislar un gen se recurre generalmente a sus productos: ARN o proteína, que servirán de sondas específicas para seleccionar dicho gen entre los muchos millares que contiene el genoma. No podemos recurrir a ese método a la hora de abordar los genes implicados en el desarrollo; por la sencilla razón de que se desconocen sus productos de antemano. Por tanto, para clonarlos hubo que aprovechar ingeniosas técnicas en las que concurrían la información genética, la información citológica y la manipulación bioquímica del ADN. Se han clonado ya, parcial o totalmente, una docena de complejos génicos involucrados en el desarrollo de *Drosophila*. Estos procedimientos de clonaje proporcionan ADN de la región donde se encuentra el complejo de interés; cumple luego identificar, dentro del ADN clonado, la zona o zonas que comprende realmente el complejo. El criterio de identificación nos lo aportan las mutaciones, que permiten comparar el ADN del insecto normal (tipo silvestre) y el del mutante. Evidentemente, si una mutación, cuyo efecto observable en el insecto (fenotipo) es el que define el complejo de interés, muestra una alteración específica en la región de ADN clonado, se puede concluir que la zona donde se encuentra esta alteración es importante para la función del complejo génico y, por definición, forma parte del mismo. Por tanto, el primer paso en la caracterización del ADN clonado consiste en identificar la naturaleza y la posición de las lesiones correspondientes a mutaciones. Algunos de los resultados más interesantes pueden resumirse tomando como ejemplos el complejo *bithorax* (C-BX), estudiado principalmente en los laboratorios de D. Hogness, de la Universidad de Stanford, y W. Bender, de la Facultad de Medicina de Harvard, ambos en Estados Unidos, y el complejo *achaete-scute* (C-AS), investigado por el grupo que encabeza el autor en el Centro de Biología Molecular de Madrid. El C-AS esta implicado en el desarrollo del sistema nervioso central del embrión y en el del sistema nervioso periférico del adulto; más específicamente, controla el número y la localización de las cerdas inervadas, las cuales constituyen unos órganos sensoriales de la mosca.

Uno de los datos interesantes es que tanto en el C-AS como en el C-BX prácticamente todas las mutaciones estudiadas presentan lesiones grandes en el ADN. Así, una parte de las mutaciones son reordenamientos cromosómicos, es decir, roturas del ADN dentro del C-AS que dividen a éste en dos partes; una y otra quedan situadas en regiones distintas del mismo cromosoma (inversiones) o en cromosomas distintos (translocaciones). Las restantes mutaciones, denominadas "puntuales", por no mostrar aberraciones cromosómicas detectables al microscopio, poseen inserciones de 0,6 a 7 kilobases de ADN extraño al C-AS o deleciones de 1,6 a 18 kilobases (1 kilobase es igual a mil pares de nucleótidos, las subunidades constituyentes de las cadenas de ADN). Por tanto, las mutaciones puntuales propiamente dichas, esto es, cambios de un nucleótido por otro, inserciones o deleciones de unos pocos nucleótidos, parecen poco frecuentes en estos complejos.

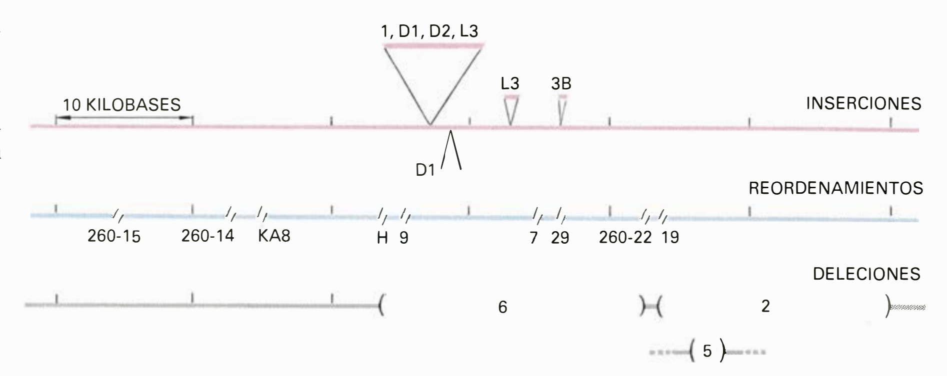
En segundo lugar, resulta sorprendente el gran tamaño de estos complejos. Vemos que en el C-AS las lesiones correspondientes a mutaciones se esparcen a lo largo de cerca de 60 kilobases y todavía quedan puntos de ruptura de reordenamientos cromosómicos más a la izquierda de la zona clonada, por lo que el C-AS puede sobrepasar fácilmente las 100 kilobases. Los datos de D. Hogness y W. Bender indican que el C-BX llega hasta las 300 kilobases. Es más, el ADN de ambos complejos, al nivel de resolución estudiado, está constituido por secuencias únicas en el genoma; es decir, no tiene repeticiones de secuencia cuantitativamente importantes dentro o fuera de los complejos. Considerando que el tamaño promedio de la mayoría de genes estudiados en la actualidad e implicados en otras funciones es de pocas kilobases, estos datos resaltan la complejidad de estos genes involucrados en la diferenciación celular. Por otro lado, el hecho de que mutaciones asociadas a lesiones del ADN muy separadas entre sí, como por ejemplo *scute*^{29} y *scute*$^{260\text{-}15}$, tengan fenotipos parecidos (supresión de quetas sensoriales), sugiere que las lesiones afectan a funciones relacionadas. Así pues, a pesar de la individualidad del ADN en cada región del complejo, hay a lo largo del mismo una similitud funcional. Estos datos, aparentemente contradictorios, comienzan a tener explicación con el estudio de los productos de transcripción de estos complejos.

En el C-BX se han detectado moléculas de ARN de unas pocas kilobases que presentan homología con dos regiones del complejo separadas entre sí por unas 70 kilobases. Esto indica que la transcripción del ADN daría lugar a un transcrito primario de por lo menos 70 kilobases, que se procesaría enzimáticamente, eliminándose grandes regiones internas del mismo sin expresión reconocida (intrones), para dar el

transcrito final de pocas kilobases. (Por transcrito primario se entiende la copia ARN del ADN, tal cual la ha sintetizado la ARN polimerasa.) Una situación análoga parece existir en el C-AS. En éste hemos detectado cuatro transcritos mayoritarios de 1,5 a 2,5 kilobases, teniendo cada uno de ellos una expresión característica a lo largo del desarrollo. Por lo menos uno parece sintetizarse a partir de un transcrito primario de gran tamaño. Es más, en el C-BX y C-AS, así como en el complejo *Notch*, estudiado por S. Artavanis-Tsakonas y M. Young, en las Universidades de Yale y Rockefeller respectivamente, hay datos que sugieren que los cambios de transcritos observados a lo largo del desarrollo podrían originarse al modificarse la forma de procesar un único o unos pocos transcritos primarios. De esta manera, ciertas regiones del ARN que en una etapa del desarrollo formarían parte de un intrón y serían eliminadas, en otra se conservarían y formarían parte de un transcrito procesado.

La existencia de transcritos primarios de gran tamaño, que abarcan regiones muy distantes del ADN, ayuda a explicar la unidad funcional observada a lo largo de estos complejos. Así, lesiones muy distantes unas de otras en la cadena de ADN pueden tener efectos parecidos al bloquear la transcripción o procesamiento de unos mismos transcritos. Ahora bien, si consideramos que prácticamente cada mutación de un complejo tiene un fenotipo característico, el bloqueo o interferencia que significa cada lesión debe ser distinta y, en consecuencia, tanto su naturaleza como su posición serán críticas a la hora de determinar de qué forma ocurre tal obstrucción. También hay que señalar que las lesiones pueden ubicarse en regiones transcritas, con lo que podrían entorpecer directamente la transcripción o el procesamiento, o en regiones no transcritas, con lo que afectarían a la transcripción a través de cambios estructurales del ADN. Parece claro que el estudio de los efectos de estas lesiones deberá proporcionar datos de gran interés para comprender la función de los complejos.

Lo antedicho revela que la caracterización molecular de los genes implicados en el desarrollo se encuentra todavía en sus principios. La investigación actual se centra en el estudio de los transcritos para definir cómo estan organizadas las regiones donde se sintetizan y cómo se lleva a cabo y se controla su expresión. En una siguiente etapa habrá que identificar las proteínas codificadas por los transcritos, localizarlas a nivel celular y estudiar sus propiedades y función. Es de esperar que estos conocimientos ayuden a entender el fascinante problema de cómo se lleva a término el programa de desarrollo codificado en el ADN de los organismos. (Juan Modolell.)



Mapa simplificado de las lesiones del ADN encontradas en mutantes del complejo génico **achaete-scute**. *Las mutaciones se han identificado por sus superíndices; por ejemplo D1 significa* **scute**^{D1}. *La incertidumbre en la situación de las lesiones oscila entre 0,1 y 3 kilobases*

Arrugas en cuatro dimensiones

Pocas ramas de la matemática han podido parecer más alejadas de la descripción del mundo físico que la topología. Aunque vemos que los físicos describen a veces características de un mundo topológicamente equivalente a un toro, la mayoría de los topólogos se encuentran más cómodos en espacios multidimensionales que en los espacios físicamente relevantes de dimensiones tres y cuatro. Mas he aquí que dos matemáticos acaban de descubrir que se puede construir al menos un espacio de dimensión cuatro cuya estructura global difiere bastante de la variedad lisamente curvada del espacio-tiempo a que se han acostumbrado los físicos.

La posibilidad de una nueva estructura para el espacio-tiempo de dimensión cuatro se planteó a raíz de los trabajos de Michael Freedman, de la Universidad de California en San Diego, quien se ha ocupado de la clasificación de las variedades topológicas de dimensión cuatro. Podemos imaginar una variedad topológica en el espacio de dimensión tres ordinario como una masa de yeso agujereada por muchos sitios. La variedad es la propia superficie de la masa; de ella se dice que está sumergida en el espacio de dimensión tres. Las variedades se clasifican, desde el punto de vista topológico, de acuerdo con el número de agujeros que aparecen cuando se las sumerge en un espacio de dimensión superior. Así, el donut y la taza de café con asa son topológicamente equivalentes por tener un único agujero cada objeto; y si la taza de café fuese de un material maleable, la podríamos deformar hasta convertirla en un donut.

El trabajo de Freedman sobre la clasificación de variedades de dimensión cuatro utiliza una noción más abstracta de agujero; sería la de un hueco en la variedad. Este tipo de agujeros topológicos subyacen a una célebre conjetura anunciada en 1904 por Poincaré. Para entender la conjetura, consideremos un lazo hecho con un trozo de hilo, que se halla sobre la superficie de una esfera, donde puede moverse con entera libertad. Si se tira de los extremos del hilo, el lazo colapsará en un punto. Pero si hacemos el mismo experimento sobre un toro, quizá resulte imposible que el lazo colapse. Hay dos maneras de disponer un lazo sobre la superficie de un toro, de suerte que no se pueda deformar en un punto. Consiste la primera en hacer que el hilo dé la vuelta al agujero y, la segunda, que lo haga a través del agujero y alrededor del "brazo" sólido del toro. Paradójicamente, el toro, que posee un único agujero, tiene dos agujeros topológicos. Poincaré conjeturó que si todo lazo de una variedad tridimensional se puede deformar en un punto, entonces la variedad es topológicamente equivalente a la "superficie tridimensional" de una esfera de dimensión cuatro.

Nadie ha demostrado todavía la conjetura de Poincaré, aunque Stephen Smale, de la Universidad de California en Berkeley, sí lo consiguiera en 1959 para una conjetura análoga referente a variedades de dimensión superior o igual a cinco. Prosiguiendo esta clasificación, Freedman ha demostrado la conjetura para variedades de dimensión cuatro. En dimensión cuatro, la conjetura consta de dos partes: no sólo cualquier lazo sumergido en la variedad ha de poder colapsar en un punto, sino que debe estarle permitido también a

cualquier esfera de dimensión dos sumergida en la misma. Así por ejemplo, en el espacio ordinario un globo puede achicarse indefinidamente sin que halle ningún obstáculo. Si una canica dentro del globo representara un agujero en el espacio, entonces el globo no puede deshincharse sin que llegue un momento en que se lo impida el agujero.

En el curso de sus investigaciones, Freedman ha construido una notable variedad de dimensión cuatro que posee ocho agujeros. En otras palabras, existen ocho maneras independientes de que una esfera sumergida en la variedad no pueda colapsarse hasta un punto. La sospecha de que existiera tal variedad se venía abrigando desde hacía muchos años, pero la demostración de su realidad no llegaría hasta el trabajo de Freedman. La variedad se denomina E^8.

Propiedad importante de las variedades topológicas es la de ser o no lisas. Del carácter "lisa" nos da una idea la aplicación de una calcomanía sobre el cristal de una ventana. Si la calcomanía se aplica con cuidado, se pegará al cristal tersa. Pero si se pega primero el perímetro, entonces aparecerán arrugas en el interior de la calcomanía cuando la presionemos sobre la superficie. Aunque se pueden desplazar las arrugas y confinarlas a un área pequeña del interior de la calcomanía, no lograremos eliminarlas a no ser que despeguemos parte del perímetro. E^8 está tan rizada que siempre contiene arrugas.

Aunque E^8 no se puede alisar, después del trabajo de Freedman quedaba abierta la cuestión de si era o no posible alisar la variedad que resultaba de pegar dos copias de E^8, una estructura que los topólogos llaman $E^8 + E^8$. Recientemente, Simon Donaldson, alumno de la Universidad de Oxford, ha demostrado que tampoco $E^8 + E^8$ puede ser alisada. Aunque Donaldson se sirvió del trabajo de Clifford Taubes, de la Universidad de Harvard, y de Karen Uhlenbeck, de la de Illinois en Chicago, su línea argumental es totalmente propia. Cuando Freedman se enteró de este resultado pudo percibir las consecuencias: había más de una manera de construir un espacio-tiempo liso y de dimensión cuatro. El espacio-tiempo ordinario, curvado o no, se denomina R^4; el nuevo espacio-tiempo, falso espacio-tiempo.

¿Qué aspecto presenta el falso R^4? Nadie lo sabe a ciencia cierta. Si hacemos caso de Robion Kirby, de Berkeley, está "convolucionado de un modo nunca visto antes". Freedman lo compara a la gráfica de la función $\mathrm{sen}(1/x)$, que cruza el eje de las x una infinidad de veces para valores de x entre 0 y 1. Un espacio-tiempo con la estructura topológica de un falso R^4 se vería como el espacio-tiempo ordinario en la vecindad del observador, pero la superficie de una esfera de dimensión cuatro lejana (definida como el conjunto de puntos equidistantes del observador en el espacio-tiempo) tendría una morfología muy compleja, por no decir inimaginable. No está claro todavía que sólo haya un candidato a falso R^4. De acuerdo con Freedman, quizás exista una familia infinita de falsos R^4. Y si los hay, la perspectiva sugiere nuevas cuestiones (similares a las suscitadas en el siglo XIX cuando se desarrollaron las geometrías no euclidianas) acerca de la posición privilegiada del R^4 ordinario en la descripción del universo que habitamos.

El elemento 109

El anuncio de la síntesis de un nuevo elemento pesado, el que constituye el número 109 de la tabla periódica, ha renovado el interés de los físicos nucleares por materializar una anhelada secuencia de la tabla, la de los elementos superpesados, cuyos números atómicos se sitúan por encima del 114. Se cree haber identificado un átomo del elemento 109 entre los productos de una colisión nuclear extraordinariamente "suave" desencadenada el verano pasado en el laboratorio de la Sociedad de Investigación de Iones Pesados de Darmstadt, Alemania Federal. Peter Armbruster y Gottfried Münzenberg dirigían el grupo de investigadores.

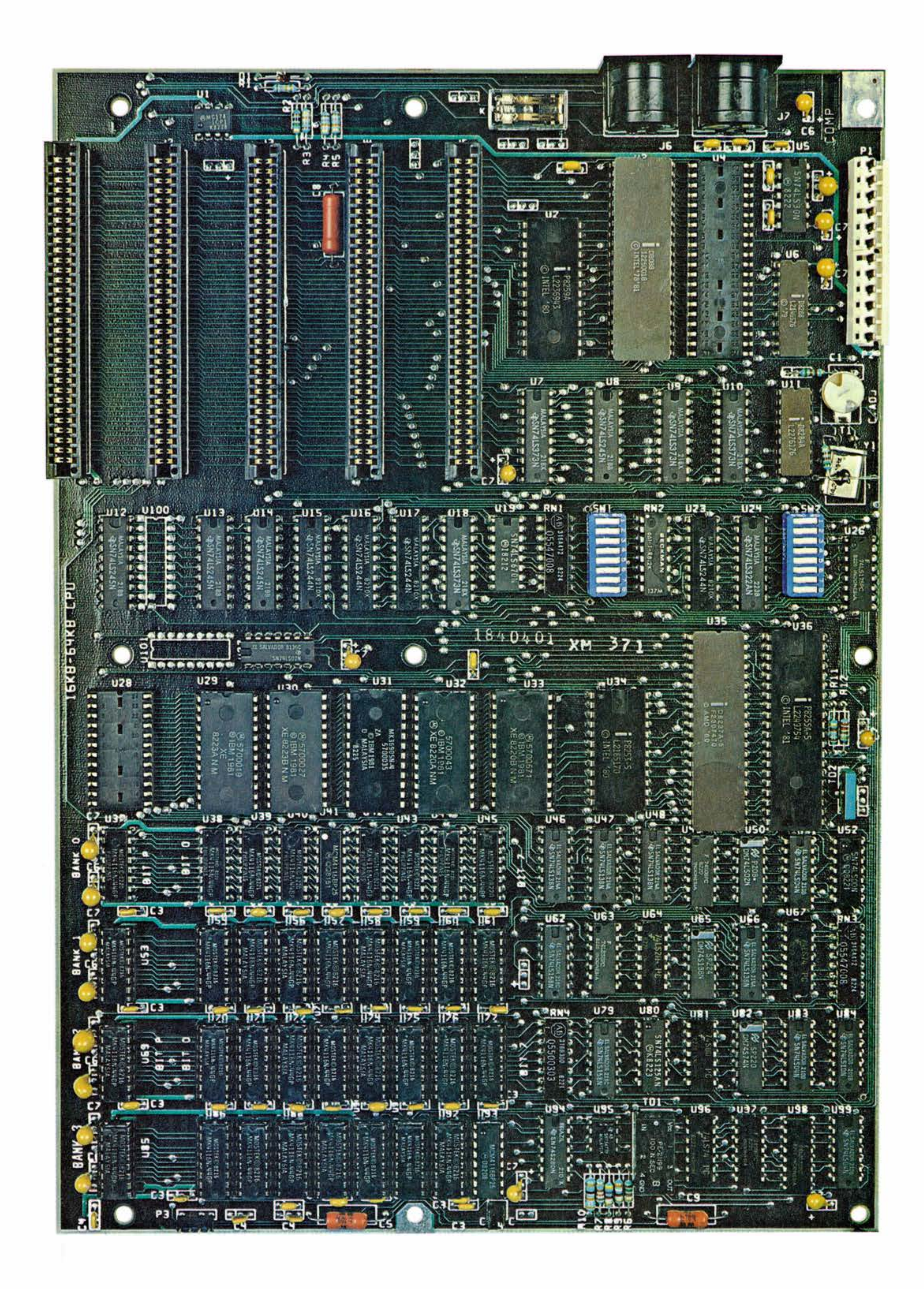
El nuevo elemento se generó en el acelerador de iones pesados del laboratorio (denominado Unilac), por medio de la técnica de fusión fría. Se aceleró un haz de iones de hierro-58, que se proyectó contra una delgada lámina de bismuto-209. En principio cabe esperar que la colisión de un núcleo de hierro de 58 nucleones (26 protones y 32 neutrones) y un núcleo de bismuto de 209 nucleones (83 protones y 126 neutrones) produzca un núcleo compuesto de 267 nucleones (109 protones y 158 neutrones), aunque sólo bajo ciertas condiciones, estrictamente definidas. La colisión debe tener suficiente energía para que se fusionen los núcleos, pero sin la violencia que los separe inmediatamente.

Según los cálculos, la fusión fría de hierro-58 con bismuto-209, aún bajo condiciones perfectamente controladas, sería un acontecimiento muy poco frecuente: un caso por cada 100 billones de colisiones nucleares. Armbruster, Münzenberg y sus colegas planearon una sesión de diez días en el acelerador de Darmstadt, a finales del pasado agosto, para detectar pruebas de la formación del elemento 109. Por fin, el último día del programa, el 29 de agosto, registró el detector un caso posible. Cuatro mediciones de las propiedades del átomo recién formado y de los productos de su desintegración confirmaron que, durante unos cinco milisegundos, había existido un núcleo del elemento 109.

La síntesis del elemento 109, caso de poderla repetir, establecería un nuevo récord nuclear: sería el elemento más pesado detectado hasta hoy. Los elementos naturales, que ocupan todos los números atómicos hasta el 92, y los artificiales, hasta el 106, se conocen desde hace ya algún tiempo. En 1981, Armbruster, Münzenberg y sus colegas identificaron el elemento 107, tras una acalorada disputa con los investigadores del Instituto Conjunto de Investigación Nuclear de Dubna. (Se supone que el elemento 108 es menos estable que el 107 y el 109.)

El aparente éxito de la técnica de fusión fría ha animado a aplicar ese método a la añeja búsqueda de elementos superpesados. La verosimilitud de que existan tales elementos se basa en el modelo de núcleo en capas, según el cual los nucleones ocupan capas sucesivas, de forma similar a las capas de electrones que determinan las propiedades químicas de los átomos. Cada capa puede albergar un número determinado de nucleones; al completarse una capa, aumenta la estabilidad. Se supone que el llenado de las capas de nucleones en los elementos cuyos números atómicos van del 114 al 126 conferirá una estabilidd poco usual a esos hipotéticos núcleos. Los elementos situados en esa "isla de estabilidad" se detectarían mucho más fácilmente que los que van del 109 al 114.

Los físicos germanos han unido ahora sus esfuerzos con un grupo del laboratorio Lawrence de Berkeley, encabezado por el veterano sintetizador de elementos pesados Albert Ghiorso, en un proyecto que pretende obtener, por fusión fría, el elemento 116. Los experimentos en curso contemplan el bombardeo de un blanco de curio-248 con un haz acelerado de iones calcio-48 en el acelerador de iones pesados SuperHILAC del laboratorio de Berkeley y en la máquina de Darmstadt. Ghiorso, quien apostó una vez a su colega Glenn Seabor 100 dólares a que nunca se encontrarían elementos superpesados, ha rebajado la probabilidad de fracaso del experimento hasta sólo 100 contra uno.


16KB-64KB CPU
1840401 XM 371
BANK 0
BANK 1
BANK 3
COMP
P1
J6
J7
SW1
SW2

Ordenadores personales

Reseña de su constitución material y lógica, de sus aplicaciones y de su proliferación en nuestros días. Por ser accesibles a personal profano, prometen traernos la revolución informática tan largamente anunciada

Hoo-min D. Toong y Amar Gupta

Si a lo largo de los 25 últimos años la industria aeronáutica hubiese experimentado la espectacular evolución que ha vivido la industria informática, un Boeing 767 costaría hoy unas 60.000 pesetas y circunvalaría el globo terrestre en 20 minutos, consumiendo unos 20 litros de combustible. A tal proeza equivaldría, en líneas generales, la reducción de costo, el aumento de velocidad y el ahorro en consumo energético de los ordenadores. El precio de los dispositivos lógicos para ordenadores está cayendo a razón de un 25 por ciento anual, y el de los dispositivos de memoria, a razón de un 40 por ciento. En 25 años, la velocidad de cómputo se ha multiplicado por 200. En ese mismo período, el consumo de energía y el tamaño de los ordenadores de capacidad de cálculo semejante han quedado divididos por 10.000.

Todo ello ha resultado en el advenimiento del ordenador personal, que por menos de 65.000 pesetas pone a disposición de cualquier individuo aproximadamente la misma potencia de cómputo de que disponían los grandes ordenadores de principios del decenio 1960-70, o un miniordenador de los primeros años 70. Hace 20 años, el costo de los ordenadores limitaba su aplicación a aquellos casos en que la inversión se justificaba si la máquina atendía las necesidades de una gran organización. Los miniordenadores introducidos hace unos diez años eran apropiados para departamentos, o para grupos de trabajo en el seno de una tal organización. Hoy, los ordenadores personales pueden formar parte de la dotación del puesto de trabajo de individuos. Además, lo mismo que ha llegado a ser económicamente factible proporcionar un ordenador al trabajador individual, también los desarrollos técnicos han ido haciendo la comunicación, o "interfase", entre hombre y máquina más y más "amigable", hasta el punto de que hoy son accesibles a personal no formado específicamente un amplio abanico de funciones computarizadas.

El primer ordenador personal salió al mercado en 1975. A fines del año pasado había, sólo en los Estados Unidos, más de un millón de ordenadores personales. En 1981, la venta de este tipo de aparatos y sus accesorios alcanzó en dicho país unos 2200 millones de dólares (alrededor de un cuarto de billón de pesetas) y se espera que en 1985 tal cifra rebase los 6000 millones de dólares. Desde que la industria electrónica logró, a fines de la década de los 50, inscribir circuitos miniaturizados en una oblea de silicio, se ha venido hablando de "revolución informática". Sin embargo, más que revolución ha habido evolución; eso sí: una evolución sostenida y notablemente rápida. Puede, no obstante, que con el advenimiento de los ordenadores personales quede abierto el camino hacia una verdadera revolución de la gestión comercial, de la forma en que la gente organiza sus asuntos personales y quién sabe si en su forma misma de razonar.

Un ordenador es, en sustancia, una

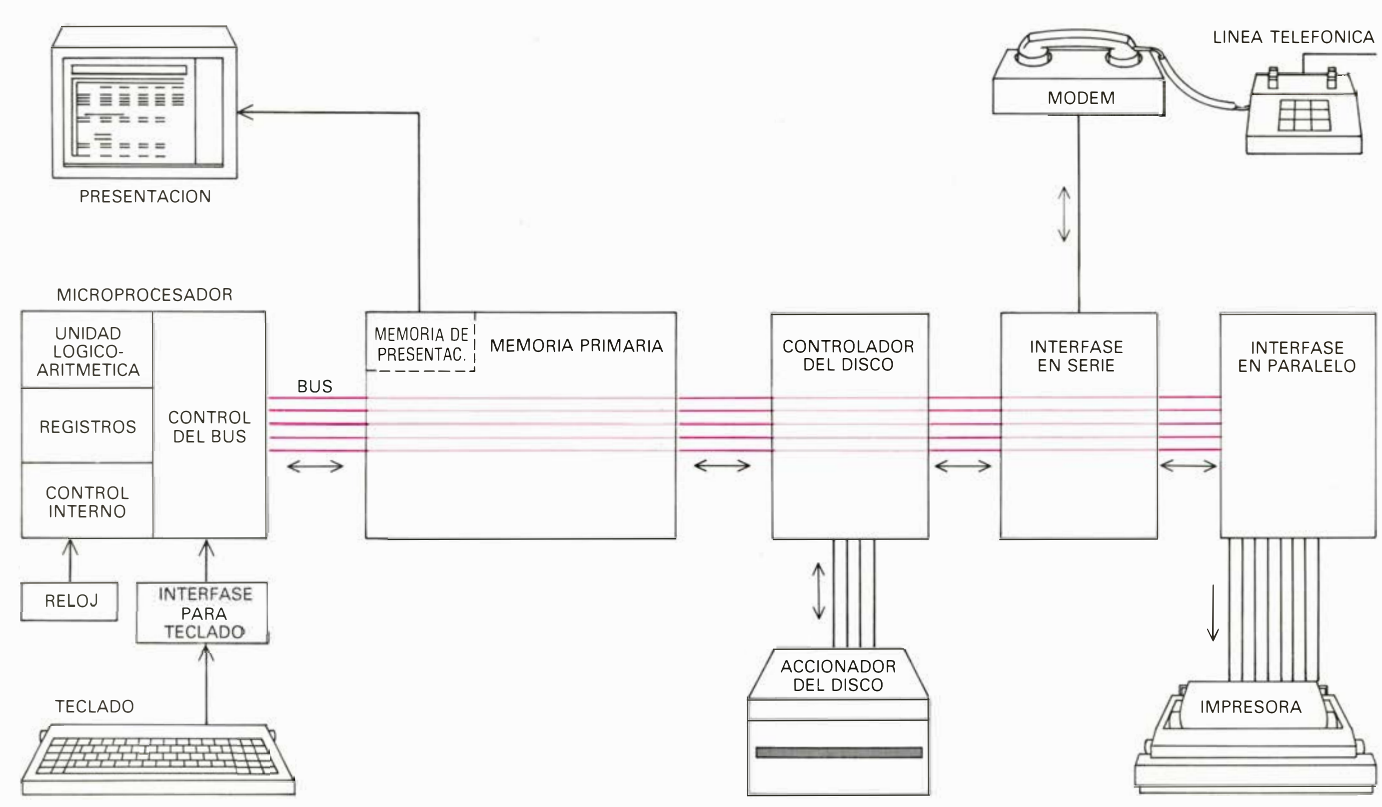
1. PLACA PRINCIPAL del circuito del Ordenador Personal de IBM, fotografiada en la página adyacente. En el dibujo de esta página se muestra el emplazamiento de sus elementos más importantes. La placa mide 216 por 305 milímetros, y a ella se encuentran fijados gran número de circuitos integrados en pastillas ("chips") de silicio. Cada pastilla es una lámina cuadrada o rectangular de unos 6 milímetros de lado, encapsulada en un bloque rectangular de plástico provisto de electrodos. Las pastillas y demás elementos componentes, como resistencias y condensadores, están interconectados mediante pistas conductoras impresas en la placa. El microprocesador (situado en la parte superior derecha), un 8088, es de 16 bits, y lo produce la Intel Corporation. Contiene alrededor de 20.000 transistores y funciona a una frecuencia de casi cinco millones de ciclos por segundo. Los "programas del sistema" están almacenados permanentemente en memoria "sólo de lectura" (ROM), mientras que la memoria "de acceso aleatorio" (RAM) sirve para almacenar programas y datos que puede alterar el usuario según su conveniencia.

máquina que recibe, almacena, manipula y comunica información. Para ello procede a descomponer la tarea a realizar en operaciones lógicas ejecutables mediante números binarios, es decir, series ordenadas de ceros y unos y a efectuar sobre estas series millares o millones de operaciones cada segundo. En el corazón del ordenador se encuentra la unidad central de proceso, encargada de ejecutar las operaciones aritméticas y lógicas fundamentales y de supervisar el funcionamiento de todo el sistema. En los ordenadores personales la unidad central es un microprocesador, esto es, un circuito integrado construido sobre una delgada lámina de silicio de muy pocos milímetros de lado, lámina que en la jerga de la microelectrónica se denomina pastilla, o "chip". Otras pastillas de silicio constituyen la memoria primaria del ordenador, en la que pueden almacenarse datos e instrucciones. Las hay que se encargan de gobernar la entrada y salida de datos; o se ocupan de las operaciones de control. Los microcircuitos se montan sobre una placa de plástico o fibra de vidrio, sólida y mecánicamente rígida, sobre cuyas caras está impresa una red de pistas conductoras que los interconecta y a través de la cual reciben la energía de funcionamiento. La placa se encierra en un gabinete; en ocasiones, en un mismo gabinete se interconectan varias de estas placas.

Anatomía de un ordenador

El ordenador recibe información a través de un teclado, o bien se le transfiere desde un sistema de almacenamiento secundario en cinta o disco magnético. La salida del ordenador puede quedar expuesta en una pantalla, ya sea el iconoscopio (llamado también tubo de rayos catódicos) del monitor de vídeo del propio ordenador, o la de un televisor doméstico ordinario. También puede imprimirse la salida en papel con auxilio de una unidad impresora acoplable eléctricamente al ordenador. Pueden acoplarse también dispositivos llamados "modems" (modulador-desmoduladores) capaces de convertir las señales generadas por el ordenador en señales aptas para su transmisión por línea telefónica; otro modem las regenera a su llegada.

Los microcircuitos y demás elementos electrónicos, juntamente con los dispositivos auxiliares, "periféricos", constituyen el soporte físico, el "hardware", del ordenador. Por sí solo, el soporte físico nada puede hacer; para funcionar requiere todo un sistema de programas, es decir, repertorios de instrucciones, que reciben el nombre de soporte informático o lógico ("software"). El corazón del soporte lógico es un "sistema operativo" que controla las operaciones del ordenador y se encarga de la gestión del flujo de información. El sistema operativo actúa de intermediario entre la máquina y el usuario humano y entre la máquina y los programas "de aplicación", que permiten al ordenador llevar a cabo tareas específicas: resolver, por ejemplo, una ecuación diferencial, calcular las partidas de una nómina, o escribir una carta. Por lo común, los programas se almacenan en dispositivos de memoria secundaria, y van siendo leídos y cargados en la me-



2. EQUIPO ELECTROMECANICO, o soporte físico ("hardware"), de un ordenador personal; cuenta con dispositivos para procesar y almacenar información y para comunicarse con el usuario o con otros dispositivos electrónicos. Un conjunto de conductores paralelos, llamado *bus* (*color*), interconecta los elementos más importantes. La unidad de proceso, que por lo común no sólo contiene la pastilla del microprocesador sino también otros microcircuitos auxiliares, ejecuta prácticamente todos los cálculos y se ocupa de controlar el conjunto del sistema. La información puede introducirse en el sistema a través de un teclado. Al pulsar una tecla, se genera una señal codificada, característica de esta tecla; el código está almacenado en la memoria de presentación de datos y, por tanto, el correspondiente signo aparecerá sobre la pantalla del iconoscopio. La memoria primaria, formada por microcircuitos de memoria construidos sobre una laminilla de material semiconductor, conserva los programas y datos que en ese momento se están utilizando; es una memoria de acceso aleatorio, lo que significa que el contenido de una cualquiera de sus celdillas puede modificarse independientemente de las demás. El almacenamiento en discos permite una capacidad de almacenamiento mayor que la memoria primaria, pero el acceso a la información que contienen es más lento y ésta ha de recuperarse en bloques mucho mayores. Los sistemas de interfase facilitan la conexión al ordenador de otros dispositivos, como impresoras y "modems" (que permiten a los ordenadores comunicarse a través de la red telefónica). En una interfase "en serie" la información se transmite a razón de un bit tras otro; en las interfases "en paralelo", un sistema múltiple transporta varios bits (ocho, en general) al mismo tiempo.

moria primaria conforme se necesitan para aplicaciones concretas.

El ordenador personal

Un ordenador personal es un pequeño ordenador basado en un microprocesador; se trata, en una palabra, de un microordenador. No todos los microordenadores son, sin embargo, ordenadores personales. Puede que un microordenador esté dedicado a una tarea específica, por ejemplo, el control de una máquina herramienta, o de la inyección de combustible en un motor de automóvil. Puede ser un procesador de textos, un vídeo-juego o una calculadora de bolsillo, máquinas todas ellas que no son exactamente ordenadores. Un ordenador personal es otra cosa: un ordenador autónomo que pone a disposición de un individuo todo un amplio abanico de posibilidades. Definiremos como ordenador personal un sistema de las siguientes características:

1. El costo de un sistema completo es inferior a 5000 dólares (unas 640.000 pesetas).
2. El sistema lleva incorporado –o puede conectarse a– una memoria secundaria de cintas o discos magnéticos.
3. El microprocesador puede soportar una memoria primaria con capacidad para al menos 64 kilobytes. (Un "kilobyte", pronúnciese "kilobáit", es igual a 2^{10}, o sea, 1024, bytes. Un "byte" es una cadena de ocho bits, es decir, ocho dígitos binarios. Un byte puede representar un carácter alfabético o uno o dos dígitos decimales. Una memoria de 64 kilobytes puede almacenar 65.536 caracteres, unas 10.000 palabras de texto español.)
4. El ordenador puede manejar al menos un lenguaje de alto nivel, como Basic, Fortran o Cobol. En lenguajes de este tipo es posible enunciar las instrucciones con un alto grado de abstracción, y sin pararse a considerar las operaciones individuales que ejecuta el soporte físico.
5. El sistema operativo facilita un diálogo interactivo; el ordenador responde inmediatamente (o cuando menos rápidamente) a las acciones y peticiones del usuario.
6. La distribución de los equipos se realiza a través de canales de comercialización en masa, y los sistemas de venta están dirigidos sobre todo a personas que no han trabajado anteriormente con ordenadores.
7. La flexibilidad del sistema le permite aceptar una amplia gama de programas capaces de atender aplicaciones muy variadas; no es, pues, un equipo diseñado con una finalidad específica, ni está concebido para una categoría exclusiva de compradores.

Sin duda, esta definición tendrá que ir cambiando conforme los avances tecnológicos hagan posible –y el mercado lo demande– que el sistema básico disponga de mayor memoria, o que soporte físico y lógico ofrezcan características más especializadas, o más potentes. Una vez definido qué es un ordenador personal –por necesidad, un tanto arbitrariamente– describiremos ahora con algún detalle sus componentes esenciales.

Microprocesador y memoria

Dos de los principales elementos que determinan la potencia de cómputo de un microprocesador son el tamaño de sus "palabras", que gobierna la "anchura" del camino recorrido por los datos, y la frecuencia de su reloj electrónico, que sincroniza las operaciones que se efectúan en el seno del ordenador. La tendencia actual apunta hacia palabras de mayor tamaño y frecuencias de reloj cada vez más elevadas. Conforme las palabras aumentan de tamaño, las operaciones pueden efectuarse con menor número de ciclos de máquina; al aumentar la frecuencia, más ciclos se ejecutan cada segundo. En general, un mayor tamaño de palabra conlleva la posibilidad de acceder a mayor volumen de memoria. La primera generación de ordenadores personales auténticos, que hicieron aparición entre 1977 y 1981, se basaba en microprocesadores de ocho bits; los sistemas más recientes se sirven de dispositivos de 16 bits. Se dispone ya de microprocesadores de 32 bits, que pronto formarán parte de sistemas de cómputo completos. Hoy las pastillas de ocho bits cuestan alrededor de las 600 pesetas, las de 16 bits rondan precios diez veces mayores y las de 32 bits están en torno a unas 32.000 pesetas. Conforme los perfeccionamientos técnicos consigan reducir sus precios irán apareciendo en el mercado más ordenadores personales con microprocesadores de 32 bits. Sin embargo, hasta aproximadamente 1985, el tamaño convencional será de 16 bits. En lo referente a frecuencia del reloj interno, se ha pasado de un megahertz (un millón de ciclos por segundo), alcanzado hace años, a los diez o más megahertz de nuestros días.

Hay dos tipos de memoria primaria: memoria ROM (read-only memory) cuyo contenido tan sólo puede ser "leído", pero no puede ser modificado, y memoria RAM (random-access memory). La memoria ROM, que también se llama memoria "muerta", sirve para conservar de manera permanente e indeleble información "cargada" en fábrica. El usuario no puede alterar su contenido. Cuando la máquina está proyectada para una sola finalidad, como puede ser un procesador de palabras, la información alojada en ROM puede incluir también el programa de aplicación. En el caso de un ordenador personal versátil, tal memoria ha de contener cuando menos los "programas del sistema" más fundamentales, esto es, los programas que ponen en marcha el aparato al pulsar el conmutador de encendido, o los encargados de interpretar la pulsación de los signos del teclado, o de hacer imprimir un archivo almacenado en el ordenador. Con el abaratamiento de las ROM se ha ido generalizando entre los fabricantes la tendencia a presentar en memoria ROM más y más programas, en lugar de darles soporte en otros sistemas de almacenamiento secundario.

La memoria de acceso aleatorio se llama también memoria de lectura/escritura, y a veces, memoria "viva". En ella pueden registrarse y leerse nuevas informaciones tantas veces cuantas sean necesarias. Las pastillas de RAM almacenan información que puede alterarse, incluidos tanto programas como datos. Por ejemplo, cuando un programa correspondiente a una determinada aplicación es "leído" por el ordenador en un disco de almacenamiento secundario y "anotado" en una RAM interna, el microprocesador puede tener acceso a las instrucciones del programa. Los microcircuitos de memoria viva (RAM) alojan información en una especie de casillero rectangular electrónico, cada una de cuyas cuadrículas tiene capacidad para un bit. La densidad de las memorias comercialmente disponibles, que podemos definir como número de bits que puede albergar un solo microcircuito, se ha visto multiplicada por 64 en los diez últimos años, y consecuentemente, el costo por bit instalado es hoy unas 50 veces menor. Hace cinco años ninguna pastilla de RAM alcanzaba a almacenar más de 16 kilobits (16.384 bits); hay ahora varios ordenadores personales provistos de microcircuitos de 64 kilobits; hacia 1984 habrá pastillas para 256 kilobits.

A pesar de que individualmente los microcircuitos de memoria sean organizaciones de bits, por lo común la transferencia de información desde y hacia la memoria primaria se efectúa en bloques de ocho bits, es decir, en bytes; lo que explica que la capacidad de memo-

ria del ordenador se mida en estas unidades. Un ordenador personal típico viene dotado de una memoria RAM comprendida entre 16 y 64 kilobytes, generalmente ampliables mediante módulos enchufables o placas de expansión. En general, es buena norma que, al adquirir un sistema, la memoria de éste sea, como mínimo, suficiente para albergar al más largo de los programas cuya ejecución se prevea. La mayoría de los paquetes de programas disponibles y a la venta dan indicación de la cantidad de memoria requerida.

El medio ordinario de almacenamiento secundario es el disco magnético flexible ("floppy disk"). Son éstos discos de mylar recubiertos por una o ambas caras de un material magnético. Los discos flexibles actuales se fabrican en diámetros de 5¼ y 8 pulgadas (13,33 y 20,32 centímetros, respectivamente). En ellos, la información queda registrada en pistas concéntricas de minúsculas regiones magnetizadas; los dígitos binarios 0 y 1 quedan representados por inversiones en la dirección de magnetización. La grabación de informaciones en el disco y su recuperación posterior se consigue gracias a una cabeza de lectura/escritura que se desplaza en sentido radial sobre el disco giratorio hasta una pista determinada. Esta pista se encuentra, a su vez, dividida en cierto número de sectores; como regla general, la información se registra o lee a razón de un sector cada vez. Según el formato que se adopte en cada caso particular, el disco contiene entre 8 y 26 sectores por pista; cada sector puede alojar entre 128 y 512 bytes de datos. La capacidad total de almacenamiento de un disco flexible varía según la densidad con que se almacenan los datos a lo largo de las pistas (que puede alcanzar hasta unos 2700 bits por centímetro), la densidad de las pistas concéntricas (hasta seis por centímetro en dirección radial) y el número de segmentos en que se divide cada pista. La mayoría de los discos flexibles actuales tienen capacidades que oscilan entre los 125 y los 500 kilobytes; se están preparando discos de capacidad aún mayor.

Una alternativa de precio muy superior a los discos flexibles son los discos Winchester; en éstos el recubrimiento magnético se aplica sobre un disco rígido de aluminio. Una unidad de disco Winchester para ordenador personal puede tener una capacidad de 5 a 50 megabytes (millones de bytes) y la transferencia de datos es mucho más rápida que en un disco flexible. Por otra parte, el disco Winchester está sellado permanentemente en la unidad que lo acciona, mientras que los discos flexibles pueden retirarse y sustituirse por otros.

Un sistema de almacenamiento secundario más sencillo y barato es la cinta magnética de audio. Una sola cinta "cassette" puede alojar tanta información como un disco flexible de capacidad mediana o baja. Pero el tiempo de acceso a una dirección determinada, esto es, a un determinado lugar de almacenamiento, es muchísimo mayor que con disco, al estar en la cinta dispuesta la información en una única sucesión lineal. Rasgo común importante de todos los sistemas de almacenamiento secundario por medios magnéticos es que la información se conserva aunque se desconecte el ordenador.

Salida

En un ordenador personal el medio de salida primario es visual, por lo común la pantalla de un iconoscopio (tubo de rayos catódicos), ya sea de un monitor incorporado al ordenador, ya del televisor doméstico del usuario. Empiezan a ser competitivos los sistemas de visualización basados en tecnología de cristal líquido o de descarga gaseosa, especialmente para sistemas pequeños o portátiles. Las imágenes y

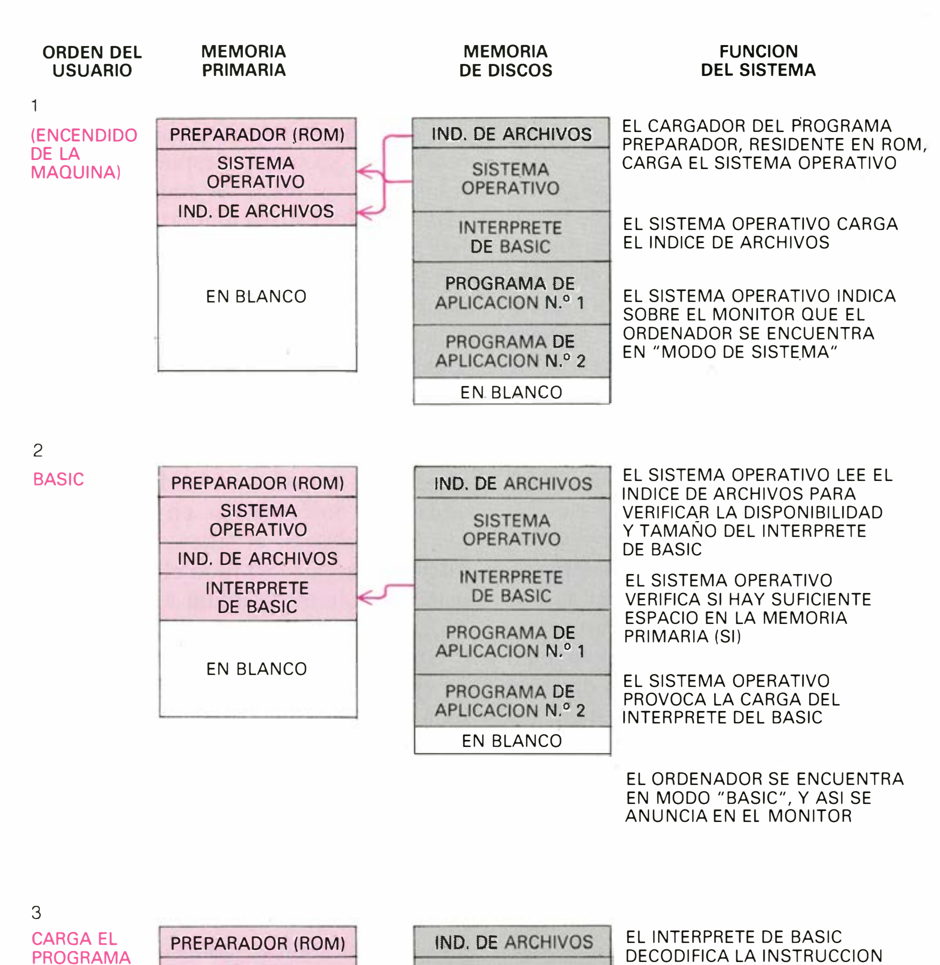
PROGRAMAS DE SERVICIO GENERAL
PREPARADOR (BOOTSTRAP)
CONCATENADOR
EDITOR
DEPURADOR
TRAZADOR
ETCETERA
COMPILADORES; INTERPRETES Y ENSAMBLADORES
LENGUAJES DE ENSAMBLADO
BASIC
FORTRAN
PASCAL
C
ADA
ETCETERA
PROGRAMAS DE APLICACION
PROCESO DE TEXTOS
IMPUESTO SOBRE LA RENTA
JUEGOS
SISTEMA DE INVENTARIO
LISTA POSTAL
GRAFICOS
ETCETERA
GESTION DE ARCHIVOS Y PLANIFICACION DE TAREAS
GESTION DE ENTRADAS Y SALIDAS
GESTION DE LA MEMORIA
EXCITADOR DE PERIFERICO
EXCITADOR DE PERIFERICO
EXCITADOR DE PERIFERICO
EXCITADOR DE PERIFERICO
EXCITADOR DE PERIFERICO
TECLADO
MONITOR
IMPRESORA
MODEM
ALMACENAMIENTO EN DISCOS
MEMORIA PRINCIPAL

3. SOPORTE LOGICO ("software") de un ordenador; está centrado en su sistema operativo (*color*), sistema de programas encargados de la administración y gestión de los recursos del ordenador, supervisión del almacenamiento de los programas y de otras informaciones y coordinación de las diversas tareas. Los programas de aplicación son los encargados de efectuar alguna función, caso de así solicitarlo el usuario. En teoría pueden diseñarse programas de aplicación capaces de funcionar sin sistema operativo, pero tales programas tendrían que contener instrucciones detalladas para asignar espacios de almacenamiento, tanto en la memoria primaria como sobre los discos, y otro tanto para hacer funcionar todos los periféricos dependientes del ordenador. Estas tareas se hallan encomendadas al sistema operativo. Para poder ejecutarlos, los programas han de estar "en lenguaje de máquina" (cuyas "palabras" son cadenas de dígitos binarios). La tarea de traducción necesaria la realizan programas llamados ensambladores, compiladores e intérpretes. Ensambladores y compiladores traducen por completo los programas antes de proceder a ejecutarlos; los intérpretes van traduciendo las instrucciones una tras otra conforme se va ejecutando el programa. Diversos programas de utilidad, que a veces se consideran parte integrante del sistema operativo, pueden asistir al programador, o hacen funcionar otros programas. Por ejemplo, los programas formadores o preparadores ("bootstrap") proporcionan las instrucciones iniciales cuando el ordenador entra en funcionamiento; los programas "trazadores" examinan el estado del sistema.

caracteres necesarios para la visualización de textos están almacenadas como pautas de puntos en una ROM especial llamada generatriz de caracteres. La claridad del texto depende del número de puntos utilizados para formar cada carácter. Un monitor típico puede presentar en pantalla 24 líneas de texto, de unos 80 caracteres cada una.

La presentación de figuras y gráficos, tanto si son diseños ingenieriles como si son gráficos estadísticos o los blancos móviles de un vídeo-juego, requiere un complejo soporte lógico y grandes cantidades de memoria. Un gráfico minucioso o una curva suave exigen imágenes de alta resolución. La resolución está determinada por el número de *pixels* (*pic*ture *el*ement*s*, elementos de imagen) a que el ordenador puede dirigirse. Una imagen en blanco y negro de 280 × 192 pixels ocupa más de 50 kilobits de capacidad de memoria RAM, mientras que una imagen de 128 × 48 tan sólo requiere alrededor de seis. Muchos ordenadores personales pueden generar imágenes en color, lo que multiplica por un factor cuatro (o aún mayor) las necesidades de memoria viva. Una imagen de alta resolución, especialmente si es en color, tan sólo puede quedar claramente presentada sobre monitores específicamente diseñados para ello.

Para muchos fines es deseable disponer de una copia impresa de la salida del ordenador. Hay cierto número de tipos de impresoras, cuyo precio oscila entre márgenes muy amplios, lo mismo que su velocidad y la calidad del texto que producen. Las más económicas son las impresoras térmicas, cuyo costo es inferior a las 65.000 pesetas; en ellas la escritura se produce quemando puntos sobre un papel especial, a razón de unos 50 caracteres por segundo. Las impresoras de matriz de puntos cuestan entre 50.000 y 200.000 pesetas, y pueden ser muy rápidas: hasta 200 caracteres por segundo. Una formación de finos alambres, cuyo número puede variar entre 5 y 18, va desplazándose a través del papel. Señales emitidas por el ordenador impulsan estas finas agujas contra una cinta entintada, dejando sobre el papel una pauta de puntos. La calidad de los caracteres así formados depende en gran medida del tamaño de la matriz de puntos disponible para cada carácter; ordinariamente esta matriz es de cinco por siete, o de siete por nueve, puntos. Con programas de control adecuados y suficiente capacidad de memoria, la impresora de matriz puntual genera imágenes gráficas en blanco y negro o en color.



4. FUNCIONES DEL SISTEMA OPERATIVO. Pueden ejemplificarse en la secuencia de acontecimientos precisos para cargar en el ordenador un programa de aplicación. El acto de encender el ordenador (*1*) pone en marcha un programa preparador ("bootstrap") que carga en la memoria el sistema operativo. El sistema operativo transfiere un índice de archivos desde la memoria de discos hasta la memoria primaria; en el índice de archivos está consignada la dirección o localización de cada programa y cada archivo de datos grabado en el disco. En respuesta a la siguiente instrucción (*2*) el sistema operativo busca en el disco el intérprete de "Basic", y, tras cerciorarse de que hay en la memoria primaria suficiente espacio para él, lo carga en ella. Se le notifica al usuario que el intérprete está listo. (Algunos ordenadores ejecutan automáticamente el paso número *2* como parte de la secuencia de puesta en funcionamiento.) Se le pide después al sistema operativo que cargue el programa de aplicación (*3*) propiamente dicho. Hecho esto, el intérprete asume el control y puede ejecutarse el programa de aplicación. La salida, o resultado del proceso, será un nuevo archivo de datos, que puede transferirse al disco para su almacenamiento.

La mayoría de las impresoras térmicas o de matriz de puntos generan textos legibles, aunque bastante burdos. Para lograr una impresión de mayor calidad tipográfica son necesarios sistemas más costosos y afines a las máquinas de escribir. Uno de tales dispositivos es la impresora de "margarita" (daisy-wheel), cuyo costo no es inferior a 95.000 pesetas, capaz de imprimir hasta 55 caracteres por segundo. La cabeza impresora consta de un cubo de rueda del que parten 96 o más radios, cada uno de los cuales porta en su extremo libre una letra u otro carácter. Conforme la rueda de margarita va desplazándose a través del papel, las señales procedentes del ordenador hacen girar la rueda hasta situar el carácter a imprimir frente a un martillo, que al

percutir sobre el brazo provoca la impresión sobre el papel.

Soporte lógico

Aunque en última instancia sea el soporte físico del ordenador lo que determine su capacidad para almacenar y procesar información, raramente tiene el usuario ocasión de manejarlo directamente. Entre el usuario y el soporte físico interviene toda una jerarquía de programas, que constituyen en su conjunto el soporte lógico del ordenador.

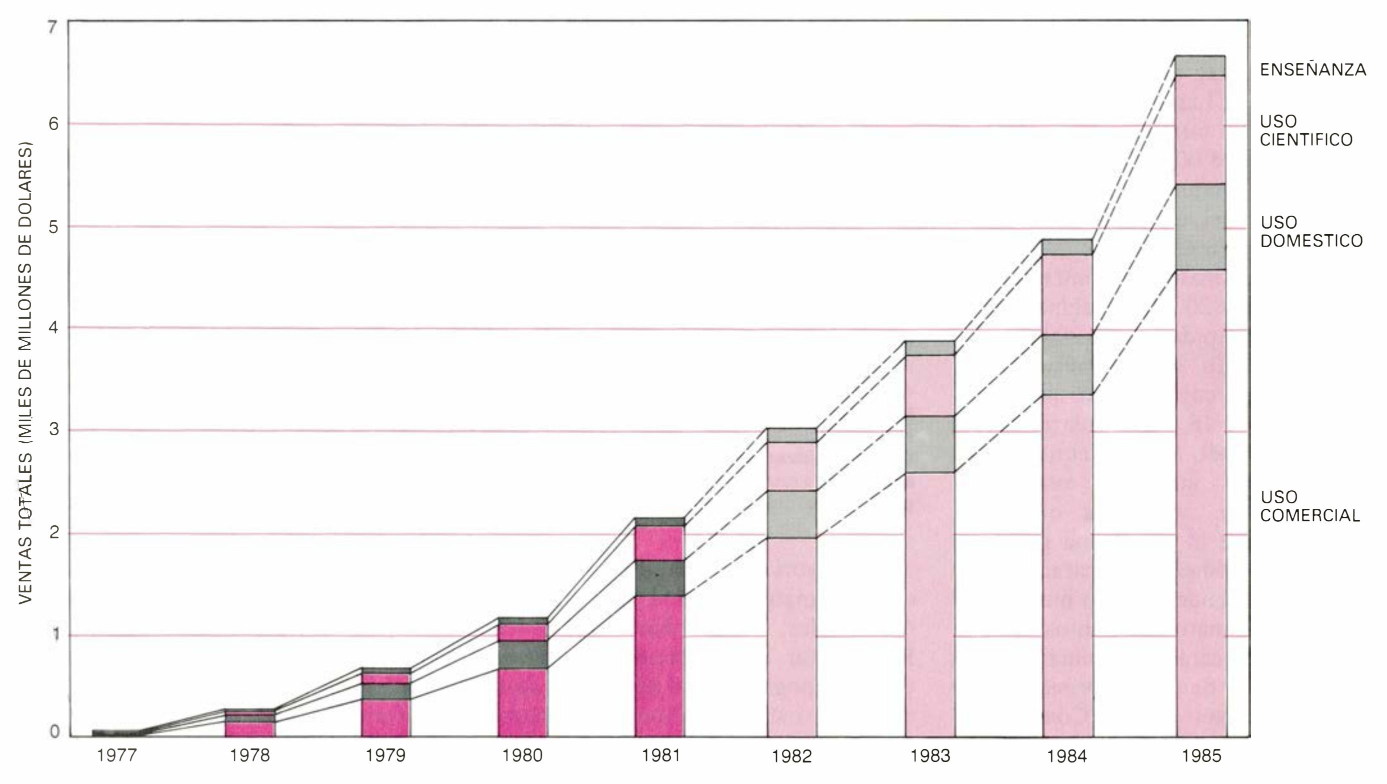
La parte del soporte lógico más íntimamente asociada con el soporte físico es el llamado sistema operativo ("operating system"). Para comprender la clase de tareas encomendadas al sistema operativo pensemos en la secuencia de operaciones a ejecutar para transferir un archivo de datos desde la memoria primaria hasta una memoria de discos magnéticos. Es necesario, ante todo, asegurarse de que hay en el disco suficiente espacio disponible para alojar la totalidad del archivo. Pudiera ser necesario borrar otros archivos anteriores para ensamblar suficiente número de sectores contiguos en blanco. Para efectuar la transferencia propiamente dicha es preciso ir extrayendo y recuperando secuencialmente porciones del archivo alojadas en la memoria primaria, combinarlas con información de control interno ("housekeeping") e ir formando bloques de datos que llenen exactamente un sector. A cada bloque de datos hay que asignarle una dirección de sector, y después hay que transmitirlo al disco. Se requiere, además, calcular ciertos números llamados "de prueba" ("checksums") que permitan la detección -y en ocasiones, la corrección- de fallos de almacenamiento o transmisión. Finalmente, es preciso registrar de alguna forma el lugar donde se ha almacenado el archivo.

Si la ejecución de todas estas tareas hubiera de hacerse bajo la supervisión directa del usuario, no valdría la pena almacenar información en ordenadores. En realidad, el sistema operativo puede encargarse de todo el proceso; el usuario humano se limita a emitir una orden, como por ejemplo, "Save file", ("Guarda el archivo"); cuando se requiere de nuevo la información contenida en él, una orden análoga, posiblemente "Load file" ("Carga el archivo"), desencadena una secuencia de acontecimientos mediante los cuales el sistema operativo recobra el archivo almacenado en disco y lo devuelve a la memoria primaria.

En casi todos los casos, los programas de aplicación se redactan pensando en que han de ser ejecutados por un determinado sistema operativo. Por otra parte, pueden existir versiones de un mismo sistema operativo adaptadas a ordenadores diferentes. Así pues, un mismo programa de aplicación podría ejecutarse en ordenadores de distintos tipos, con tal de que todos ellos utilizaran idéntico sistema operativo; en la práctica pueden requerirse algunas modificaciones.

El microprocesador del ordenador tan sólo reconoce un reducido repertorio de instrucciones, cada una de las cuales tiene que serle presentada en forma de pauta de dígitos binarios. Una de estas pautas, por ejemplo, pudiera ordenarle al procesador que cargase en un registro interno llamado acumulador un valor tomado de la memoria primaria; otra de tales pautas pudiera decirle a la máquina que sumase dos números ya presentes en el acumulador. Es posible escribir programas en este "lenguaje de máquina", pero es proceso tedioso y en el que verosímilmente se cometerán errores.

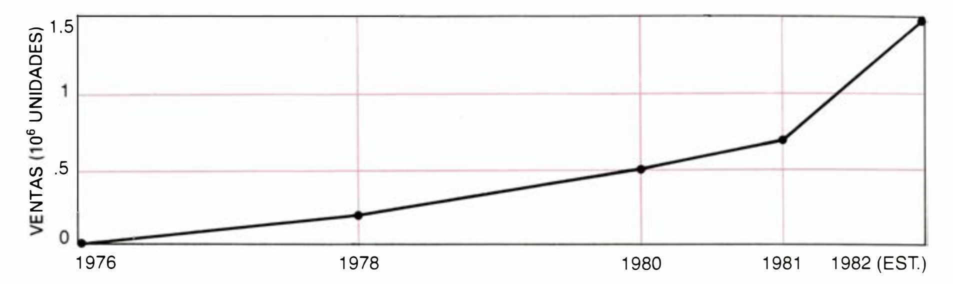
El nivel de abstracción inmediatamente superior lo forman los llamados lenguajes "de ensamblado", en los cuales las cadenas de dígitos binarios se reemplazan por símbolos y palabras más fáciles de recordar. La instrucción que ordena cargar el acumulador pudiera estar representada por LOADA, y la instrucción de sumar los contenidos del acumulador pudiera ser, sencillamente, ADD. Un programa llamado ensamblador ("assembler") reconoce cada una de estas instrucciones mnemóni-

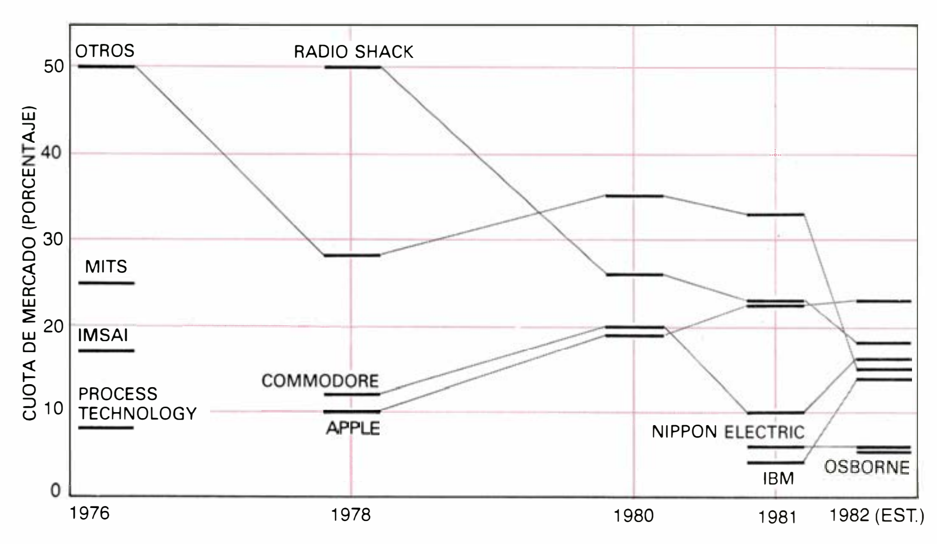


5. EL MERCADO DE ORDENADORES PERSONALES ha crecido exponencialmente durante los últimos años, y se espera que continúe con igual tendencia. Las barras del gráfico expresan en dólares el valor de las ventas de ordenadores personales de cada año en los cuatro segmentos del mercado.

cas y la traduce en la correspondiente sucesión ordenada de dígitos binarios. En algunos lenguajes de ensamblado es posible definir y apelar por un solo nombre a secuencias completas de instrucciones. Empero, los programas escritos en lenguaje ensamblador han de ir detallando todavía cada una de las operaciones que el procesador debe ejecutar; además, el programador puede tener también que llevar control del lugar reservado en la máquina a cada instrucción y cada unidad de datos.

Los lenguajes de alto nivel descargan al programador de la molestia de tener que adaptar cada algoritmo al sistema concreto de instrucciones del procesador y de tener que prestar atención a la configuración interna del soporte físico. Si hay que sumar dos cantidades, éstas pueden ser llamadas X e Y, sencillamente. En vez de tener que decirle al procesador en qué lugar de la memoria primaria se encuentran los valores a sumar, el programador especifica la propia operación, quizás en la forma $X + Y$. El programa, que ha ido llevando registro de dónde localizar las variables mencionadas, genera una secuencia de instrucciones en lenguaje de máquina que hace que sus valores sean cargados en el acumulador, y una vez allí, sumados.





6. CIFRA DE VENTAS anuales en el campo de los ordenadores personales. Se ha multiplicado por 100 en seis años; en 1981 se duplicó con creces. Las compañías precursoras de este ramo no han logrado sobrevivir, viéndose desplazadas por otras cuyos productos han interesado a un mercado más amplio.

Hay dos grandes categorías de programas, llamados intérpretes y compiladores, que traducen, a código de máquina, programas escritos a más alto nivel. Los programas escritos en lenguaje interpretado se almacenan en forma de secuencia de instrucciones de alto nivel. Cuando se pasa el programa, un segundo programa (el intérprete propiamente dicho) va por turno traduciendo cada instrucción en la serie adecuada de instrucciones de lenguaje de máquina. En cuanto la instrucción ha quedado traducida, se ejecuta inmediatamente. En el caso de un compilador, antes de comenzar a ejecutar el programa se hace la traducción completa de todas sus instrucciones. Los intérpretes tienen la ventaja de que puede observarse el resultado de cada operación individual. Por otra parte, la ejecución de los programas suele ser mucho más rápida con lenguajes compilados, pues la traducción a lenguaje de máquina se ha efectuado ya.

Uno de los primeros lenguajes de alto nivel fue el "Fortran", del que hoy existen varias versiones (llamadas "dialectos"). Los programas Fortran son compilados. Las principales aplicaciones del Fortran se encuentran en el campo científico y matemático. Para ordenadores personales, el lenguaje de alto nivel más difundido es el "Basic", puesto a punto en la década de 1960 por los especialistas del Darmouth College. La primitiva finalidad del Basic fue servir de lenguaje de introducción al Fortran para estudiantes de programación de ordenadores; pero se utiliza ya para todo tipo de aplicaciones. La mayoría de las versiones de Basic son interpretadas. Hay varias docenas más de otros lenguajes de alto nivel ejecutables en microordenador. La elección de lenguaje en que expresar un programa determinado viene muchas veces inspirada por la naturaleza del problema a considerar; así, por ejemplo, muchos investigadores en inteligencia artificial son partidarios del lenguaje llamado "Lisp". También pesan al programar consideraciones de estilo personal. El lenguaje "Pascal" ha ido ganando popularidad en estos últimos años, porque se dice que estimula y facilita la redacción de programas cuya estructura subyacente es clara y fácilmente comprensible.

Programas de aplicación

Son los programas de aplicación los que a fin de cuentas determinan la medida en que un ordenador alcanza a satisfacer las necesidades de los humanos. Por esta razón es verosímil que el poseedor de un ordenador personal termine invirtiendo más en soporte lógico que en soporte físico. Tal inversión puede hacerse, ya comprando programas, ya dedicando a su redacción y puesta a punto el nada despreciable tiempo necesario. A menos que se esté dispuesto a una intensa labor de programación, será necesario atender cuidadosamente a la amplitud de la base informática del sistema (número de aplicaciones que puede soportar) y la profundidad de la misma (número de programas disponibles para cada aplicación) antes de decidirse por la elección de uno determinado.

A la sombra del microordenador ha nacido una pujante industria productora de programas de aplicación. Muchos programas tienen finalidad muy específica. Así, hay programas para cumplimentar la declaración de renta; los hay capaces de analizar (en conjunción con el instrumental de laboratorio necesario) miles de muestras sanguíneas por hora; los hay especiales para diseñar puentes. Otros programas ofrecen, en cambio, aplicaciones más generales. Tenemos un ejemplo escogido en los programas de procesado de textos, que facilitan la redacción, corrección y resumido de todo tipo de documentos,

desde cartas y memorias informativas hasta artículos de revista, como éste.

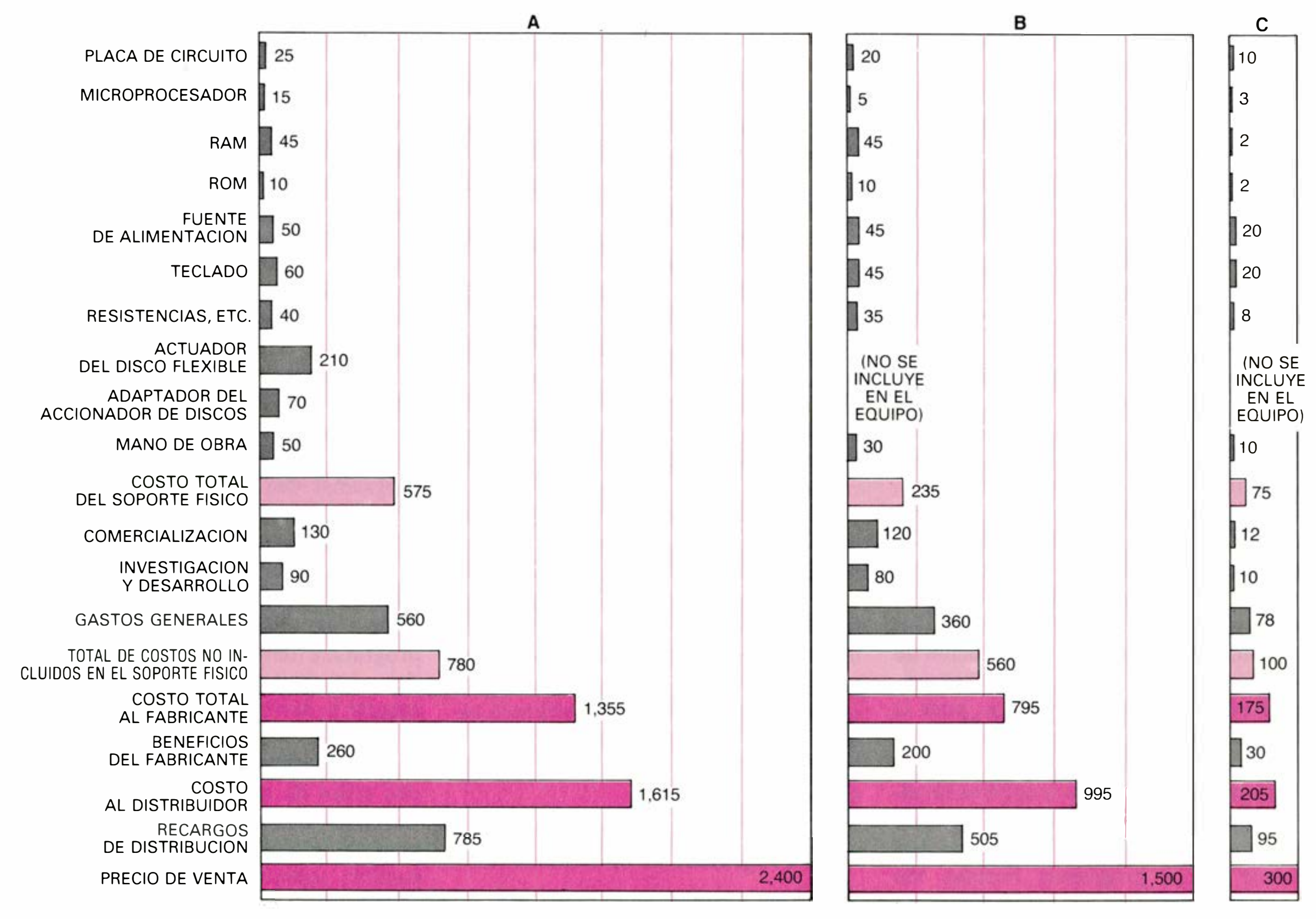
El más difundido de los programas autónomos para ordenador personal es VisiCalc, distribuido por VisiCorp. Es una "hoja de trabajo electrónica". El programa se expande en la memoria del ordenador, y hace mostrar en su pantalla una tabla de 63 columnas de anchura y 254 filas de profundidad. El usuario va "desenrollando" la hoja, haciéndola desfilar hacia arriba o abajo, y hacia la derecha o la izquierda, con el fin de dejar a la vista distintas porciones de ella. Cada posición (es decir, cada intersección de fila y columna) corresponde a un registro de la memoria. El usuario construye su propia matriz asignándole a cada registro una etiqueta, una entrada de datos o una fórmula; la correspondiente posición de la pantalla mostrará la etiqueta asignada, el dato ingresado o el resultado de aplicar la fórmula.

Tomemos un ejemplo sencillo. El contador-jefe de una compañía podría asignar la etiqueta *Caja* en el registro correspondiente a la columna *B*, fila 1 (posición *B*1), *Reservas* en *C*1 y *Total* en *D*1. Seguidamente podría asignar 30 millones a *B*2, 50 millones a *C*2 y la fórmula + *B*2 + *C*2 a la posición *D*2. La pantalla mostrará entonces 80 millones en *D*2. Si el contador cambia el dato *B*2, reduciéndolo a 20 millones, el programa reducirá el total exhibido en *D*2, dejándolo en 70 millones. Además, no es preciso que a *B*2 y *C*2 les sean asignados datos primarios, sino que éstos pueden ser funciones de datos registrados en otras posiciones.

La industria

La evolución del pequeño ordenador personal fue consecuencia, quién sabe si inevitable, del advenimiento del microprocesador. Fue en 1971 cuando la Intel Corporation logró inscribir todos los elementos de una unidad central de proceso en una sola pastilla de circuito integrado. Aquel primer microprocesador admitía palabras de tan sólo cuatro bits, aunque en el plazo de un año Intel produjo un microprocesador de ocho bits, y en 1974 comercializó una versión perfeccionada, el Intel 8080. Pronto, pequeñas compañías combinaron el 8080 con pastillas de memoria y otros componentes, produciendo así los primeros microordenadores programables para control de procesos industriales y otras aplicaciones específicas análogas. En 1975, MITS Inc. puso a punto un diseño lo bastante flexible para considerarlo el primer ordenador personal comercialmente disponible. Fue bautizado Altair 8800, y el sistema de base se vendía (fundamentalmente a entusiastas aficionados) al precio de 395 dólares despiezado, y de 621 dólares montado y en orden de operación. En aquella época, el miniordenador más económico rondaba los 6000 dólares.

El Altair ya no se fabrica. La verdad es que una de las ironías de la industria de ordenadores personales, que en sólo seis años ha visto cómo se multiplicaba por 100 su cifra de ventas, es que las firmas pioneras, como MITS, la IMSAI Manufacturing Corporation y la Processor Technology Corporation, no consiguieran traspasar la fase inicial. Sus productos eran adquiridos casi exclusivamente por aficionados y entusiastas, personas profundamente curiosas por los ordenadores, y que en la mayoría

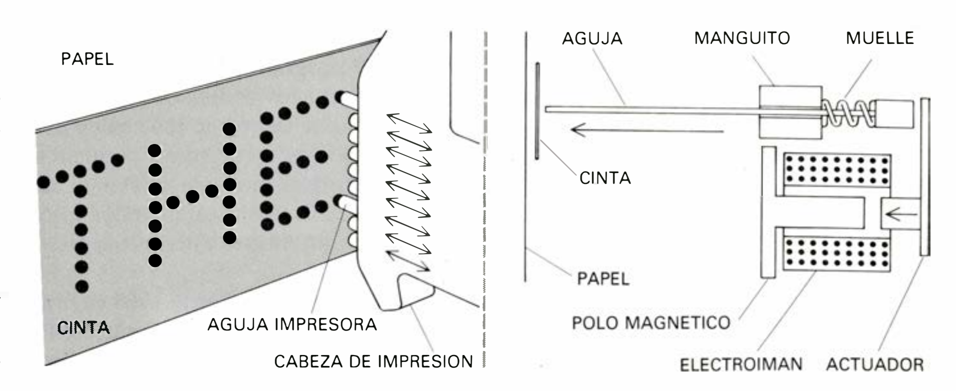


7. EL PRECIO DE VENTA de un ordenador personal refleja los costos de materiales, mano de obra y otros gastos que debe afrontar el fabricante, más los beneficios del fabricante y el margen comercial del distribuidor. Se han estimado aquí (en dólares) los costos acumulativos correspondientes a tres categorías de aparatos: un ordenador personal de precio y características relativamente elevadas (*A*), un modelo medio (*B*) y otro económico (*C*).

de los casos poseían conocimientos previos de electrónica, dispuestos, cuando no ansiosos, a luchar cuerpo a cuerpo con el soporte físico. Las compañías que desplazaron a las pioneras, logrando hacerse con el mercado hacia 1978, fueron Radio Shack, Commodore Business Machines y Apple Computer, Inc. Vieron el potencial de un más amplio mercado, hogareño o comercial, y ofrecieron sistemas "listos para enchufar", más accesibles a personas sin preparación en informática. El éxito de las compañías de "segunda generación" ha sido el aldabonazo que ha despertado a las grandes compañías de material informático, como la International Business Machines Corporation (IBM) y la Burroughs Corporation, y a los fabricantes de miniordenadores, como Digital Equipment Corporation o Hewlett-Packard Company, haciéndoles descubrir que sus mercados tradicionales pudieran verse erosionados por los ordenadores personales. Las compañías tradicionales entraron entonces en liza. Nuevas compañías, por otra parte, sienten la atracción de esta industria.

El mercado de los ordenadores personales puede dividirse en cuatro campos: comercial, doméstico, científico y educativo. El segmento comercial es ya, con mucho, el más importante. En 1981 supuso ventas de 385.000 unidades (55 por ciento del total) por un valor aproximado de 180.000 millones de pesetas, equivalentes al 64 por ciento del valor de las ventas. Sólo en los Estados Unidos hay 14 millones de empresas, cada una de las cuales, por minúscula que sea, es potencial adquiriente de un ordenador personal. Tal vez más importancia tenga todavía que en Norteamerica haya alrededor de 36 millones de empleados administrativos, y un elevado porcentaje de ellos probablemente acabe sirviéndose en su puesto de trabajo de algún tipo de pequeño ordenador.

Hoy por hoy, el ordenador personal se adapta especialmente a las necesidades de compañías pequeñas y de profesionales por cuenta propia, tales como abogados y médicos. Empero, organizaciones de mayor tamaño van lentamente haciéndose a la idea de que sus puestos de trabajo puedan ir dotados de ordenadores individuales, conectables entre sí y a una gran instalación central (dotada de grandes unidades de memoria y de potentes equipos de impresión, por ejemplo) formando redes locales de zona [véase "Mecanización del trabajo de oficina", por Vincent E. Giulinao; INVESTIGACIÓN Y CIENCIA, noviembre de 1982]. Los ordenadores personales tienen ya potencia suficiente para gestionar la mayoría de las tareas propias del puesto de trabajo, y se está procediendo a desarrollar los sistemas de redes. Se estima que hacia 1985 muchas organizaciones comerciales tendrán en funcionamiento redes de ordenadores personales.



8. IMPRESORA DE MATRIZ de puntos, económica, rápida (hasta 200 caracteres por segundo) y de utilización flexible. Puede generar caracteres estrechos, anchos o en negrita, e incluso figuras y gráficos, de acuerdo con las instrucciones que recibe del ordenador. Su cabeza impresora comprende una formación vertical de finas agujas que van excitándose selectivamente conforme el cabezal de escritura recorre transversalmente el papel; oprimen contra él una cinta entintada, formando así una configuración de puntos (*izquierda*). Cada letra mayúscula es aquí un subconjunto de una matriz que tiene siete puntos de altura y cinco de anchura; se dispone de dos puntos verticales más para los trazos descendentes de algunas letras minúsculas, como la *p*. Las agujas las van excitando solenoides individuales (*derecha*). El mecanismo representado corresponde a una impresora matricial fabricada por la compañía Epson America, Inc.

El segmento de ordenadores para uso doméstico, sin duda el más visible y objeto de mayor esfuerzo publicitario, representó en 1981 ventas de 175.000 unidades, con un valor de 45.000 millones de pesetas. La mayoría de estos aparatos fueron adquiridos con fines recreativos (fundamentalmente vídeo-juegos) aunque sirven también de ayuda a la enseñanza de niños, como procesadores de textos, centros electrónicos de mensajes e instrumentos de gestión de las finanzas familiares. El soporte lógico en fase de desarrollo permitirá una amplia variedad de nuevas aplicaciones en el hogar. Se espera que el costo medio de un equipo doméstico completo descienda desde el cuarto de millón de pesetas actual a unas 130.000 hacia 1985, y 90.000 hacia 1990.

El segmento científico del mercado supuso ventas de 105.000 unidades en 1981, con un valor de 43.000 millones de pesetas. Los ordenadores destinados a aplicaciones científicas o técnicas tienden a ser más potentes que otros ordenadores personales, y a disponer de elementos que faciliten su acoplamiento a instrumentos de medición o análisis. El correspondiente mercado se caracteriza, pues, por ofrecer productos de soporte físico especializado y con un abanico de programas igualmente especializado.

Aunque el segmento educativo sea, en potencia, muy grande, padece una crítica dependencia de la consignación de partidas presupuestarias; la asignación de fondos a la enseñanza pública suele hacerse siempre con parsimonia. No obstante, en 1981 las instituciones educativas de los Estados Unidos adquirieron 35.000 ordenadores personales, por valor total de 12.500 millones de pesetas. La instrucción con ayuda de ordenador permite al alumno una vivaz interacción con el tema de estudio en prácticamente todos los campos, y permite además que el individuo avance a su propio paso. La aptitud para trabajar con un ordenador está comenzando a ser considerada como habilidad básica necesaria, e incluso para muchas ocupaciones puede ser pronto exigida cierta capacitación de programador. Como es obvio, el lugar idóneo para adquirir la formación básica necesaria es la escuela elemental y de grado medio. Razonando que los alumnos adiestrados con ordenadores de una marca determinada se inclinarán seguramente por adquirir en su momento ordenadores de esa misma marca. Commodore ha venido ofreciendo a escuelas y facultades universitarias tres ordenadores por el precio de dos, y Apple ha propuesto donar ordenadores personales a las escuelas primarias y secundarias norteamericanas.

Principales fabricantes

A la cabeza de esta industria (de acuerdo con las estimaciones de ventas para 1982) se encuentran Apple, Radio Shack, Commodore e IBM. Aunque todas ellas están concentrando su empeño en la captación de una parte del segmento empresarial tan grande como les sea posible, se esfuerzan igualmente

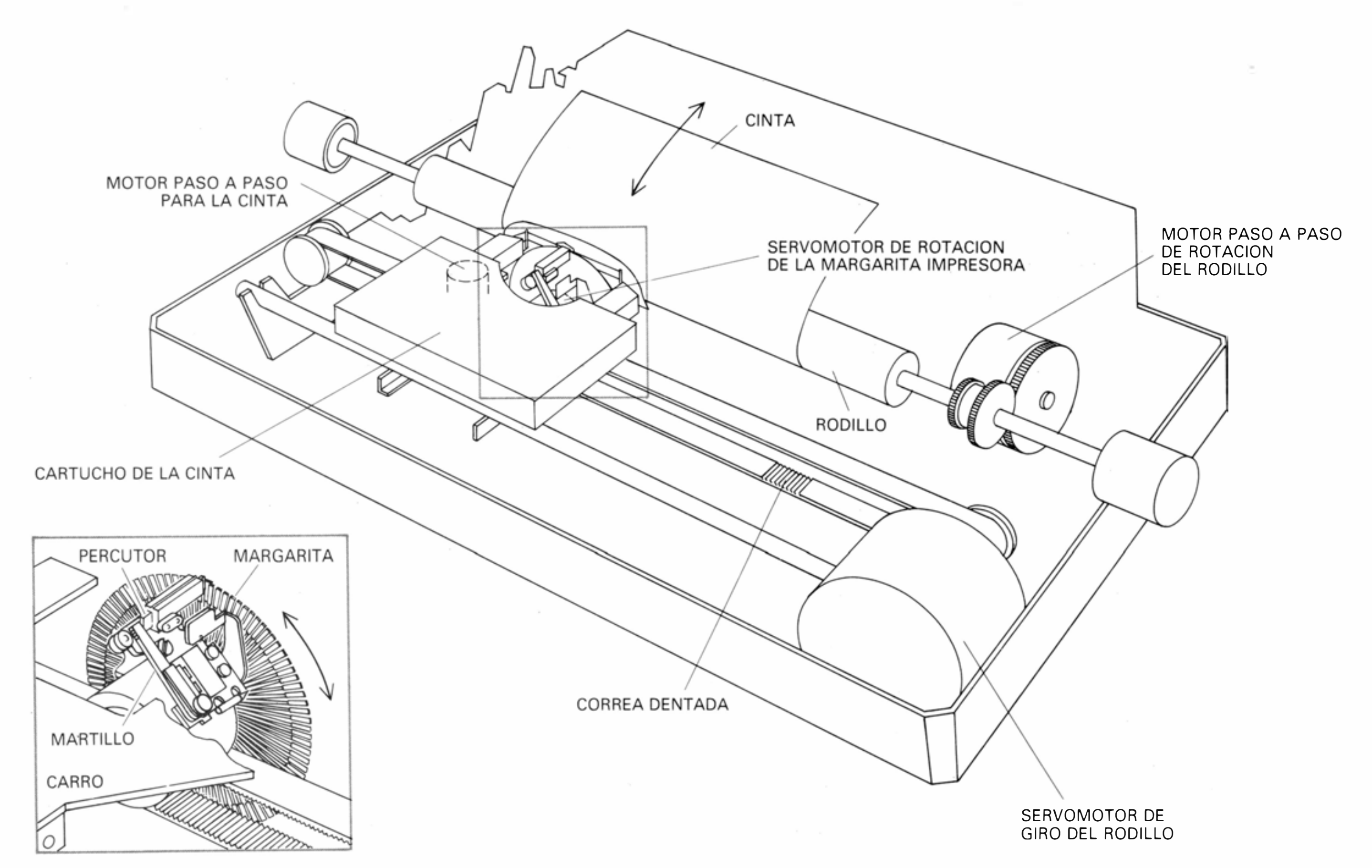
en atender a los otros segmentos del mercado.

La compañía de máximas ventas (no sólo en los Estados Unidos sino en todo el mundo) es Apple, cuyo primer prototipo se construyó en un garaje en 1976. Los cuatro primeros años de la compañía se financiaron con inversiones de particulares y capital de empresa, y aunque la compañía se convirtió en sociedad anónima en 1980, todavía el 64 por ciento de sus acciones pertenecen a gente de la casa. La cifra de ventas de Apple alcanzó en 1981 los 335 millones de dólares (casi 43.000 millones de pesetas), 2,9 veces más que el año anterior; sus beneficios fueron de 39,4 millones (unos 5000 millones de pesetas) multiplicando por 3,4 los de 1980. Apple afirma controlar el 23 por ciento del mercado norteamericano, y debe recordarse que sus ventas en él significaron solamente el 76 por ciento de su total. Gran parte del éxito de Apple se atribuye a la política comercial de la compañía, consistente en animar e inducir a otros fabricantes y distribuidores a poner a punto materiales y soporte lógico compatibles con ordenadores Apple. Hay, por ejemplo, más de 11.000 programas de aplicación disponibles para los ordenadores de Apple, un 95 por ciento de los cuales han sido creados por empresas desligadas de ella. Los tres modelos de Apple actualmente a la venta están todos ellos basados en un mismo microprocesador de ocho bits.

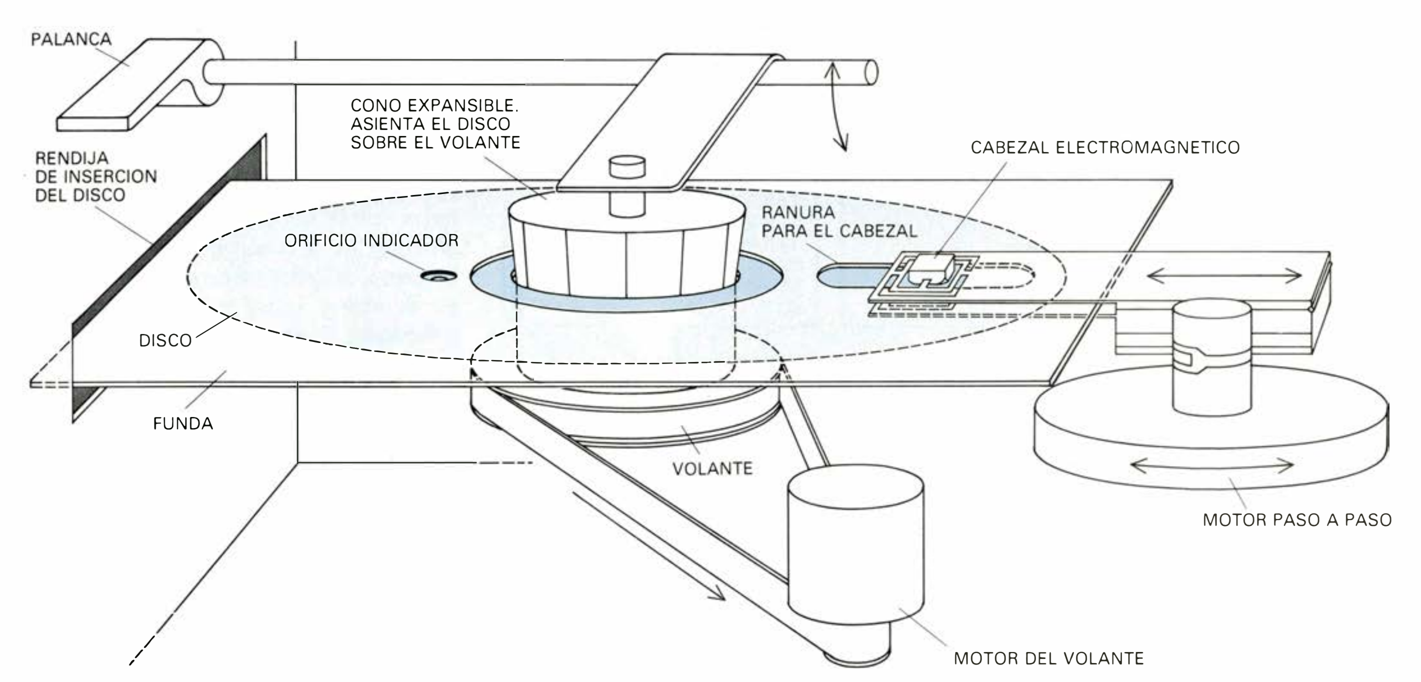
Radio Shack, que desde 1963 es propiedad de la Tandy Corporation, fabricaba y distribuía productos electrónicos mucho antes de entrar en el campo de los ordenadores, que hoy representa alrededor de la quinta parte de su cifra global de ventas. Aunque éstas han ido creciendo de forma regular, su control sobre el mercado ha disminuido, desde el 50 por ciento que ocupaba en 1978 hasta el 22 por ciento previsto para el año pasado. Radio Shack ofrece una amplia línea de productos para ordenador, muchos de ellos de fabricación propia, y dispone de una red de distribución excepcionalmente buena, pues además de sus 8000 puntos de venta de electrónica general posee una red internacional de centros especializados en ordenadores, encargados de ventas, servicio técnico, alquiler y formación del cliente. El soporte lógico de la compañía se ha desarrollado tanto en la misma compañía como en empresas independientes.

Commodore es una firma canadiense que nació en 1958 adscrita al ramo de las máquinas de escribir. En 1976 adquirió MOS Technology, empresa primitivamente fabricante del microprocesador utilizado todavía en los ordenadores de Apple y Atari. Commodore es la compañía con máxima cifra de ventas fuera de los Estados Unidos y controla el 65 por ciento del mercado europeo. Dispone de una extensa línea de productos de costo reducido (incluido un modelo mínimo que se vende por unas 19.000 pesetas) y ha conseguido buenos resultados en el terreno escolar.

IBM, la mayor de las proveedoras de equipos de proceso de datos del mundo, ha dominado desde hace mucho el mercado de los grandes ordenadores, pero no ha tenido igual éxito con los ordenadores de tamaño pequeño hasta su presentación en el campo de los ordenadores personales, a mediados de 1981. Ha conseguido hacerse con ex-



9. IMPRESORA DE "MARGARITA", que produce escritos con calidad "de imprenta", a razón de 20 a 55 caracteres por segundo. Vemos en la figura la representación esquemática de una impresora de la Qume Corporation. La rueda impresora tiene un cubo de plástico en torno al cual están dispuestos 96 radios (130 en algunos modelos). En el extremo libre de cada uno está moldeada una letra, una cifra o cualquier otro signo. En respuesta a señales del ordenador la rueda gira, bien en sentido horario, bien en sentido antihorario, hasta situar el símbolo apropiado en posición. Una vez éste allí, la rueda se detiene. Un martillo golpea contra el extremo de un percutor deslizante (con mucha más fuerza, por ejemplo, en símbolos de área grande, como una *M*, que en los pequeños, como una coma), que ataca el extremo del brazo radial y hace que éste comprima la cinta entintada contra el papel. Al tiempo que el carro y la cinta avanzan un paso, la rueda gira, situando en posición el símbolo siguiente. Las impresoras de margarita suelen ser bastante costosas.



10. LOS SISTEMAS DE DISCO FLEXIBLE registran grandes cantidades de información sobre un disco de plástico flexible cubierto con una capa de material ferromagnético. El disco gira a 300 revoluciones por minuto dentro de una funda plástica lubricada. Una cabeza electromagnética se desplaza transversalmente, sobre la superficie del disco, impulsada en el sentido de un radio por un motor paso a paso, hasta quedar situada sobre una de las pistas concéntricas donde se encuentran almacenados los datos en forma de una serie de inversiones de la dirección de magnetización. La cabeza puede leer y escribir; es decir, captar estas inversiones y recuperar la información, o imponer magnetizaciones nuevas y almacenar información. Una señal índice, cuyo paso detecta un dispositivo fotoeléctrico, sincroniza la grabación o lectura con la rotación del disco. La figura muestra esquemáticamente un sistema de arrastre para disco de doble cara, construido por Qume. Dispone de dos cabezas montadas cada una sobre suspensión cardán, capaces de registrar y leer información sobre las dos caras de un disco de 5¼ pulgadas. Sobre cada cara del disco pueden almacenarse unos 160 kilobytes de información.

traordinaria rapidez con una porción muy notable (14 por ciento) del mercado. Su estrategia ha consistido en apoyarse fuertemente en fuentes ajenas a la empresa, no sólo para el soporte lógico, la distribución y el mantenimiento de los equipos, sino también para la fabricación de éstos. Así, el accionador de los discos del ordenador personal de IBM es de la Tandon Corporation; el monitor de vídeo procede de Taiwan y la impresora es japonesa. El teclado lo pone IBM –lo mismo que la marca. IBM ha creado una casa editorial que solicita nuevos programas de soporte lógico redactados por autores ajenos a la empresa.

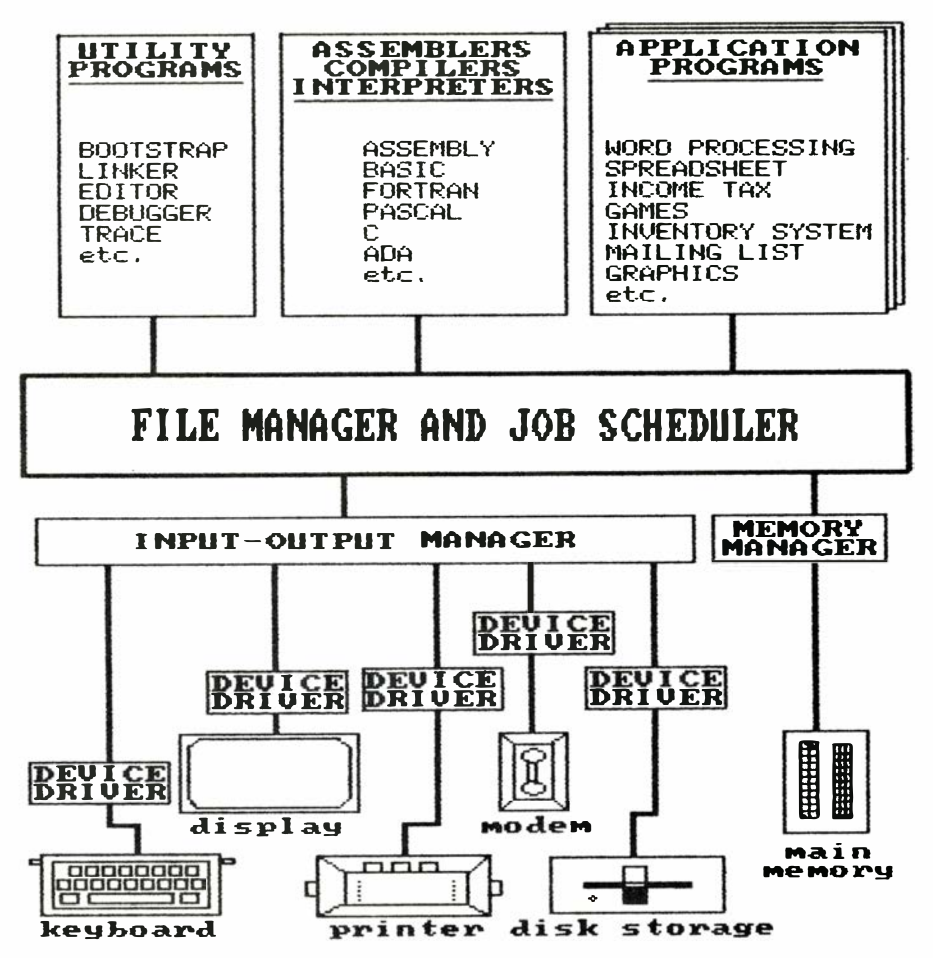
El éxito logrado por IBM comporta interesantes implicaciones para el futuro del negocio de ordenadores personales. Es ésta una industria volátil. Hay sociedades anónimas norteamericanas, como la Xerox Corporation y Atari Inc., así como cierto número de fabricantes japoneses (entre quienes destaca Nippon Electric Co., Ltd.) que se encuentran bien situadas para desplazar del mercado a los líderes de hoy. Por los flancos atacan nuevos competidores. Para evaluar cuál pueda ser su porvenir es forzoso examinar cuáles son los requisitos para el éxito comercial. A juzgar por la estrategia seguida por IBM, no parece imprescindible disponer de capacidad de fabricación propia consolidada. Más bien parece que los requisitos fundamentales sean capacidad financiera para adquirir los componentes necesarios y capacidad para comercializar con éxito un producto, distribuyéndolo rápidamente sobre una extensa área. Muchas organizaciones, entre ellas varias cuya actividad actual nada tiene que ver con la electrónica, disponen de tales capacidades, y bien podrían hacerse con la pericia técnica necesaria. Organizaciones tan dispares como CBS y Coca-Cola, Time-Life y Prudential Insurance Co. disponen de recursos financieros y vías de distribución que podrían permitirles entrar pronto en concurrencia en el mercado del ordenador personal.

Distribución

De la venta de los grandes ordenadores se encarga el personal de plantilla de la empresa fabricante, que trata directamente con el individuo o la organización interesados en utilizar el equipo. El margen de beneficio de un ordenador personal no consiente sostener una fuerza de venta directa de este tipo. Se han desarrollado cierto número de otros canales de distribución, algunos creados por los fabricantes y, otros, por empresarios detallistas.

Los minoristas independientes que gestionan una sola tienda han sufrido con los ordenadores personales una dura experiencia. Tan sólo pueden encargar reducidas cantidades de un producto, y están en general deficientemente capitalizados para poder competir con vigor; se ven por ello desplazados por cadenas de venta con franquicia, tales como Computerland, que vendió en 1981 ordenadores y accesorios por valor de 200 millones de dólares (25.600 millones de pesetas). Tales cadenas de tiendas ofrecen los productos de cierto número de fabricantes, pueden permitirse disponer de personal técnico que aconseje al comprador y proporcionan además servicio técnico y mantenimiento a largo plazo. Otras cadenas de distribución menos especializadas, dedicadas a productos electrónicos de consumo (equipos de sonido, etcétera), han incluido en sus inventarios y catálogos los ordenadores personales, si bien acusan en general la falta de experiencia en ordenadores. Conforme el soporte físico de éstos vaya haciéndose más y más fiable, y conforme vaya convencionalizándose el soporte lógico y entren en el mercado las compañías japonesas (fuertemente ligadas a las cadenas de electrónica de consumo), estas cadenas llegarán seguramente a ser uno de los principales canales de distribución.

Los grandes almacenes no han tenido, en general, mucho éxito en la venta



11. SOPORTE LOGICO para "gráficos comerciales"; aumenta en gran medida la utilidad de los ordenadores personales, y a él debe atribuirse la aceptación cada vez más amplia de esas máquinas. Esta figura, versión inglesa computarizada de la figura 3, se generó mediante ordenador personal con el programa Execu-Vision, distribuido por Digital Systems Associates, y se imprimió con el dispositivo de matriz. La tarea se efectuó en unos 15 minutos. La rápida preparación de este tipo de gráficos, e incluso de tablas, con equipos de precio bajo, es de la mayor utilidad para la elaboración de informes y presentaciones.

de ordenadores personales. Para empezar, no les es posible ofrecer mantenimiento a largo plazo. Además, un estudio ha hecho ver que el comprador de un ordenador personal hace por término medio cuatro visitas a tiendas del ramo, que consumen en total unas siete horas; los grandes almacenes no están habituados a semejantes esfuerzos de venta. Por otra parte, las tiendas especializadas en material de oficina disponen de contactos en el mercado empresarial local y pueden proporcionar a éste los equipos y el servicio técnico necesarios. Recientemente, Sears ha abierto en las grandes ciudades tiendas especializadas, dedicadas exclusivamente a ordenadores personales, equipos de tratamiento de textos y equipos auxiliares.

Los propios fabricantes patrocinan diversas vías de salida para sus productos. Radio Shack descansa principalmente en su propia cadena de almacenes detallistas. IBM, Xerox y Digital Equipment están abriendo sus propios puntos de venta con vistas a complementar sus otros canales de distribución. Texas Instruments mantiene salas de exposición con muestrarios de sus equipos, donde el cliente puede inspeccionar el material y hacer encargos que luego se cumplimentan desde un almacén central. En ciertos casos, a los fabricantes puede estarles justificado sostener una plantilla de personal de venta directa, pensando en ventas al por mayor a organizaciones oficiales, grandes empresas o instituciones de enseñanza. Radio Shack e IBM han creado plantillas de este tipo, pero corren el riesgo de suscitar antagonismos con cadenas minoristas que podrían en otro caso competir para lograr estas ventas grandes.

Las compañías de venta por correo han contado siempre con una participación importante en el mercado de ordenadores personales. Manejan partidas grandes, y ello les permite ofrecer rebajas importantes. En contrapartida, no pueden garantizar el mantenimiento de los equipos ni el servicio técnico a largo plazo. Es, además, poco probable que los detallistas que sí pueden ofrecer estos servicios vendan productos ampliamente disponibles a precios con descuento.

Una nueva vía de comercialización, característica de los ordenadores personales, es la llamada "tienda de valor añadido", que compra materiales al fabricante, adquiere o desarrolla equipos periféricos y soporte lógico para una aplicación determinada o un determinado tipo de usuario y ofrece después equipos listos para su uso. Los servicios de tal tipo de establecimientos pueden presentar particular interés a organizaciones con poca o ninguna pericia en la programación o manejo de ordenadores.

¿A quién puede servir?

A pesar de las connotaciones de la palabra "personal" y de la imagen publicitaria que muestra a la familia apiñada en torno al ordenador para hacer los deberes escolares, llevar al día las cuentas de la casa o derribar naves espaciales, resulta obvio que la mayoría de los ordenadores personales son adquiridos por organizaciones, sean comerciales o no. No necesariamente es por ello menos personal el ordenador; puede seguir estando dedicado a las necesidades de un solo individuo. Más de la quinta parte de la población laboral norteamericana se ocupa de tareas burocráticas. En muchas empresas los gastos burocráticos representan más de la mitad de los gastos totales, y estos gastos de oficina están creciendo a una tasa superior al 7 por ciento anual. Los ordenadores personales pueden incrementar la productividad de la oficina y de sus funcionarios. En organizaciones que dispongan ya de un gran ordenador, los modelos personales pueden aligerar la carga que gravita sobre la instalación central, la cual puede así dedicar más tiempo a tareas de procesado "por tandas", tales como nóminas o control de inventarios. Los ordenadores personales permiten mecanizar una amplia gama de tareas burocráticas hasta ahora efectuadas por máquinas de escribir, calculadoras de sobremesa y fotocopiadoras.

Se dice que gerentes y directores de empresa dedican más del 80 por ciento de su tiempo a preparar y asistir a reuniones y "presentaciones", a recoger información o tomar decisiones basadas en análisis de alternativas. Los ordenadores personales inciden en estas tres actividades. Nuevos programas de "gráficos comerciales" generan rápida-

mente diapositivas y material impreso para reuniones. Los discos Winchester y los programas para almacenar y gestionar grandes bases de datos ayudan al individuo a examinar una gran masa de información, discernir tendencias en ella y detectar posibles problemas. Programas de manipulación de datos, como el VisiCalc, permiten al director evaluar distintas vías de actuación y plantearse el tipo de pregunta que empieza: "¿Qué sucedería si...?" y obtener casi instantáneamente la respuesta. Tales tareas pueden en teoría llevarse a cabo con una gran instalación central de proceso de datos, pero se ejecutan con más eficacia mediante ordenadores personales, con inversión muy inferior y por individuos sin formación en informática.

Una vez dicho todo esto, sigue siendo cierto que a menudo resulta imposible predecir el papel exacto que desempeñarán en una organización los ordenadores personales. Muchas organizaciones han descubierto que más que satisfacer una necesidad concreta y conocida, la presencia de un ordenador personal sirve muchas veces para detectar una necesidad no reconocida hasta entonces como tal (lo mismo que la presencia de un médico puede sacar a la luz un problema sanitario que estaba pasando desapercibido) y actuar en consecuencia.

Es difícil saber si un individuo necesita un ordenador personal, o si podrá sacarle provecho, o simplemente disfrutar de él. Qué duda cabe de que para ciertos profesionales saltan a la vista las ventajas de tener siempre a mano un ordenador. Otras personas tal vez lleguen a comprarlo fundamentalmente porque está a su alcance, dejando para después definir sus aplicaciones. De las capacidades del ordenador emanarán las aplicaciones específicas. Un ordenador lleva el control de cosas y permite clasificarlas. Un ordenador calcula. Puede controlar una gran masa de datos, o cambiar el valor de una variable y ver qué sucede. Puede efectivamente llevar actualizada una cuenta corriente (o mejor dicho, el poseedor puede llevar actualizada la cuenta corriente con auxilio de su ordenador), puede servir de agenda, o puede conectarse a un sistema de seguridad doméstico. Por sí sola, ninguna de estas aplicaciones justifica la adquisición de un ordenador. Sin embargo, con curiosidad e ingenio, el poseedor de un ordenador personal llegará a definir aplicaciones propias, conformando el sistema según su personalidad y gusto.

Muestras de Vía Láctea

Los rayos cósmicos son núcleos procedentes de otras regiones de nuestra galaxia. Las observaciones de tales núcleos revelan que la composición isotópica de esas regiones difiere de la composición del sistema solar

Richard A. Mewaldt, Edward C. Stone y Mark E. Wiedenbeck

Los adelantos obtenidos en la tecnología e instrumentación espaciales nos permiten determinar ya la composición de muestras de materia procedente de otros puntos de la Vía Láctea, nuestra galaxia. Aunque recoger muestras de materia procedente de fuera del sistema solar trasciende nuestras posibilidades, nos es dado llegar al mismo resultado interceptando auténticos "mensajeros estelares": los rayos cósmicos. Son éstos fragmentos de materia de alta energía que proceden de otras estrellas. Los análisis de los rayos cósmicos a lo largo de los últimos cuatro años, mediante instrumentos instalados a bordo de un satélite observatorio especializado, revelan que la composición de la materia de los rayos cósmicos producidos en nuestra galaxia difiere nítidamente de la composición del Sol y otros cuerpos del sistema solar. Es probable que tales diferencias surjan de las distintas condiciones bajo las que se sintetizan los elementos en el interior profundo de las estrellas.

Aproximadamente, el 90 por ciento de la materia de nuestra galaxia es hidrógeno. El resto está constituido, en fracción muy notable, por helio; alrededor de un 1 por ciento lo constituyen elementos más pesados. A la conversión termonuclear del hidrógeno en helio y del helio en elementos más pesados débese, sobre todo, el brillo de las estrellas. La composición de una galaxia evoluciona con el tiempo a medida que algunos de los núcleos atómicos sintetizados en el interior de estrellas de gran masa salen expulsados, en las explosiones de supernova, hacia el gas interestelar, que así va acumulando gradualmente átomos de muchas estrellas. El sistema solar representa una muestra de la mezcla atómica que hace unos cinco mil millones de años se condensó del gas interestelar en una diminuta región de nuestra galaxia. No hay razones para pensar que esa muestra sea arquetipo de la galaxia entera.

Algunos de los núcleos expulsados por una explosión de supernova pueden ser acelerados por el estallido hasta casi la velocidad de la luz y convertirse en rayos cósmicos. Otros núcleos del gas interestelar pueden acelerarse hasta altas velocidades por las ondas de choque de supernovas ulteriores. Al estar cargados eléctricamente, los núcleos atómicos interaccionan con el campo magnético de la galaxia y quedan confinados durante períodos del orden de 10 millones de años antes de escapar, por fin, al espacio intergaláctico. En su transcurso por la galaxia, algunos rayos cósmicos pasan próximos a la Tierra, donde se pueden capturar y analizar. La captura y análisis, sin embargo, deben realizarse en el espacio, porque los núcleos de rayos cósmicos que penetran la atmósfera terrestre se rompen en fragmentos por choques con los núcleos de los átomos de la atmósfera antes de que puedan alcanzar la superficie de nuestro planeta.

Las primitivas investigaciones llevadas a cabo con instrumentos transportados por globos de gran altitud y por vehículos espaciales dejaron establecido que aproximadamente el uno por ciento de los rayos cósmicos eran núcleos de elementos más pesados que el hidrógeno y el helio. Esas investigaciones demostraron también que los elementos más pesados que son corrientes en los rayos cósmicos (carbono, oxígeno, magnesio, silicio y hierro) venían a presentar la misma abundancia relativa con que se encuentran en el sistema solar. Mejorando la precisión de las mediciones de rayos cósmicos, sin embargo, podemos investigar no sólo las abundancias de los núcleos atómicos, sino también las abundancias de sus isótopos, o especies de núcleos atómicos.

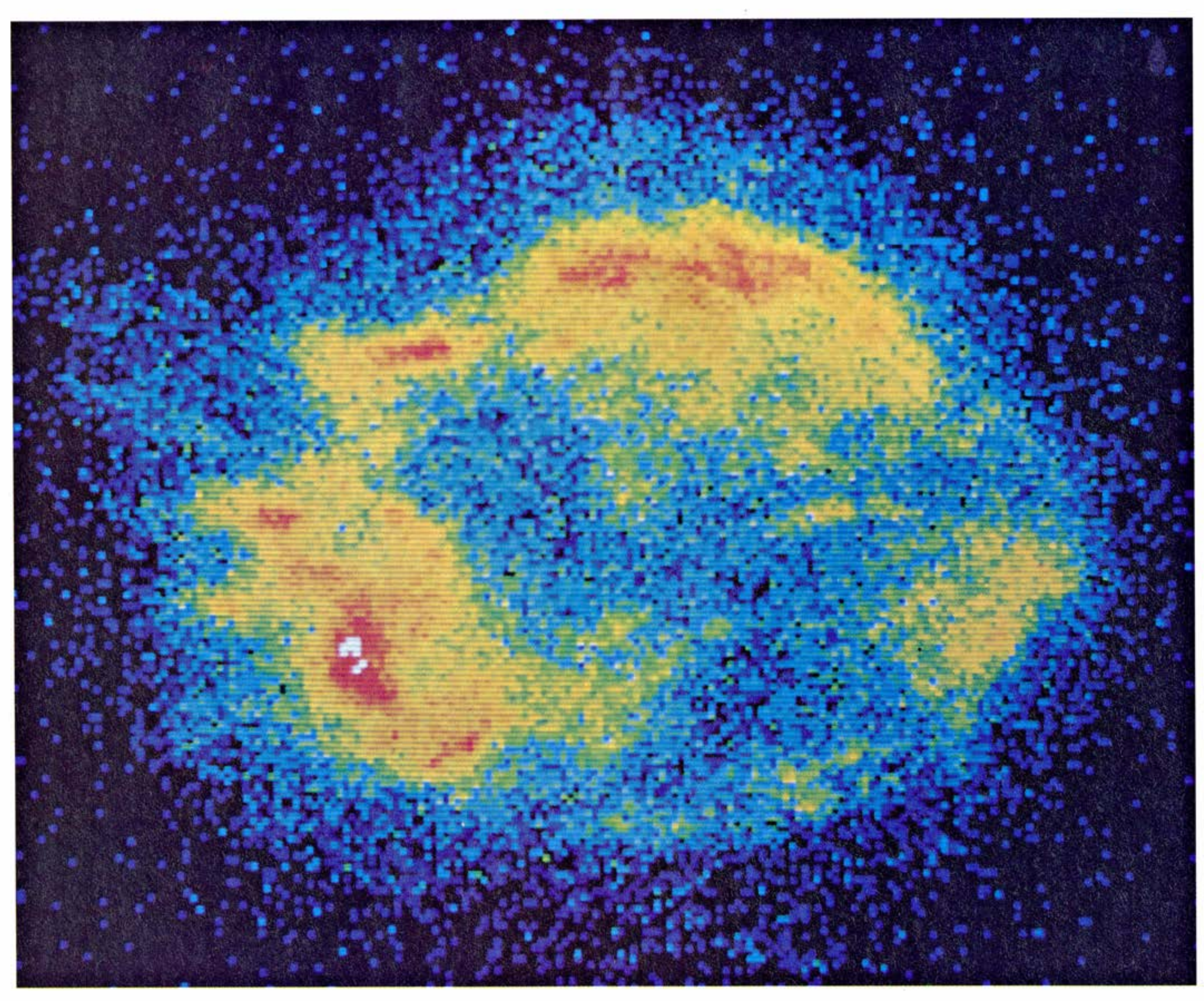
Los isótopos de los elementos se distinguen por razón de su masa. Los elementos difieren entre sí por el número de protones que encierran sus núcleos (designado por Z), y que varía desde el único protón del hidrógeno hasta los 92 protones del uranio. Los isótopos de un mismo elemento difieren entre sí por el número de neutrones que encierran sus núcleos: ninguno en el hidrógeno ordinario (H-1), uno en el hidrógeno pesado o deuterio (H-2), y así sucesivamente hasta 143 y 146 neutrones en los dos isótopos más abundantes del uranio, el U-235 y el U-238. Como los protones y neutrones tienen casi exactamente la misma masa, los isótopos de un elemento difieren entre sí en números enteros de unidades de masa atómica. Midiendo las abundancias de isótopos en los núcleos de rayos cósmicos, detectaremos materia que quizá tenga una composición distinta de la que ofrece el sistema solar y, por consiguiente, diferente historia. No sólo eso: algo averiguaremos también sobre los procesos de formación de elementos que ocurren en el interior de estrellas de tipos diferentes.

La de medir la masa de los isótopos es una operación que se realiza sin mayores dificultades en el laboratorio con un espectrómetro de masas, instrumento que tiene aproximadamente el tamaño y peso de un piano de media cola. No lo es tanto realizar la misma medición en el espacio con un instrumento que pese menos de 10 kilos y consuma una potencia inferior a 10 watt. Además, el instrumento espacial debe trabajar con núcleos individuales de alta energía y no con una voluminosa muestra de materia. Todavía más, ha de estar equipado para suministrar medidas de alta precisión, al tiempo que funcione sin mantenimiento varios años. Se construyeron dos ingenios de ésos: uno en la Universidad de California en Berkeley y otro en el Instituto de Tecnología de California. Los puso en órbita, en agosto de 1978, el Tercer Explorador Solar Internacional (ISEE-3) de la Administración Nacional de Ae-

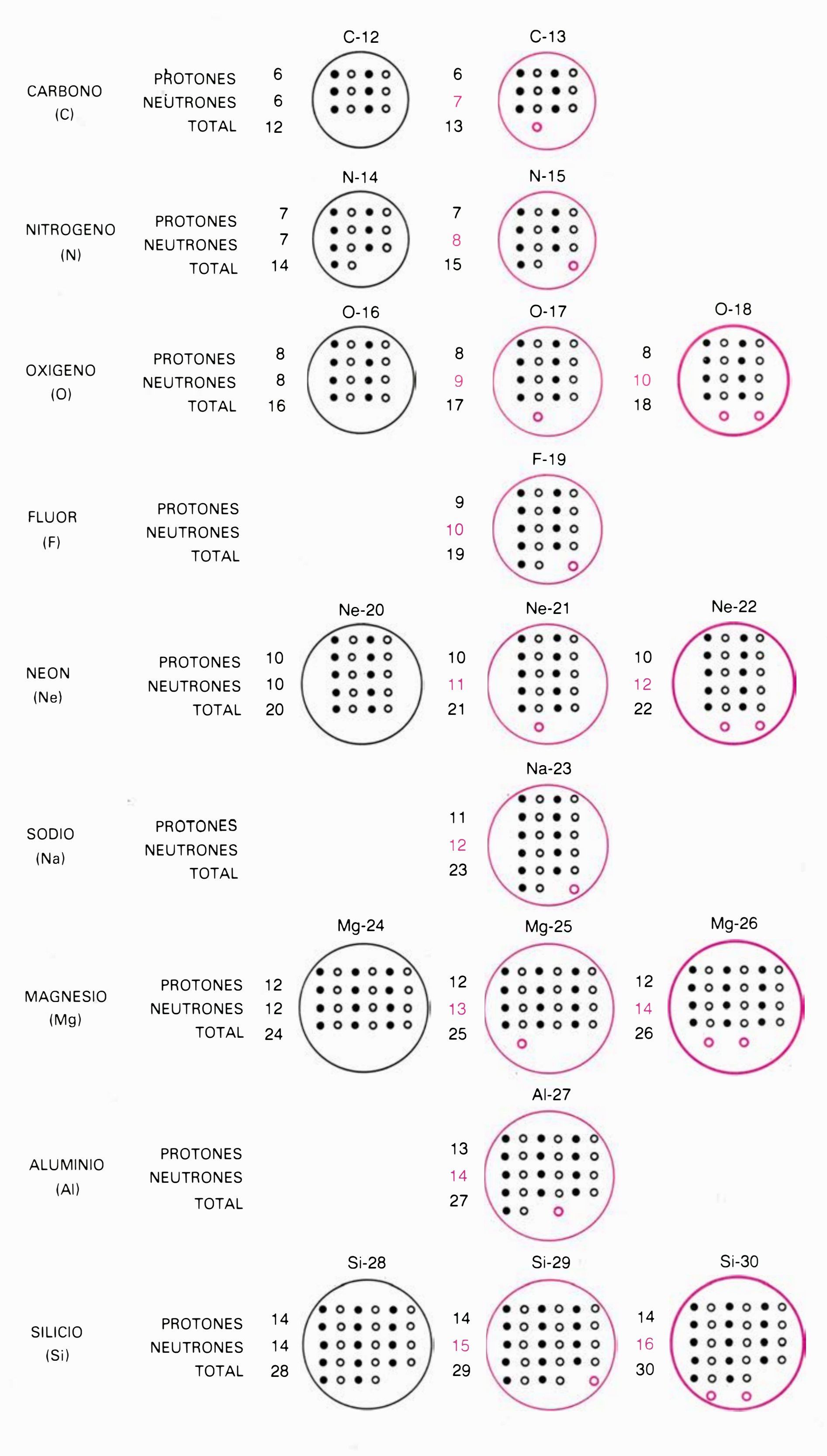
ronáutica y del Espacio, siendo el primer satélite que llevaba instrumentos proyectados específicamente para medir la composición isotópica de los núcleos pesados de los rayos cósmicos. El vehículo espacial también transporta otros diez dispositivos experimentales para medir otros fenómenos astrofísicos e interplanetarios.

El ISEE-3 se colocó en una órbita poco corriente, como un "halo" terrestre, a 1,5 millones de kilómetros de la Tierra (una centésima parte de la distancia que media entre ésta y el Sol), cerca del "punto de Lagrange", donde está sometido a una atracción gravitatoria equiparable por parte del Sol y por parte de la Tierra. Con la ayuda de pequeños impulsos de gas, el ISEE-3 sigue una trayectoria suavemente sinusoidal, que lo sitúa, de un modo alternante y a muy escasa distancia, por encima y por debajo del plano de la órbita terrestre. Cuando cruza el plano orbital, se encuentra sucesivamente un poco por delante o un poco por detrás de la línea entre la Tierra y el Sol. Así, visto desde la Tierra, el ISEE-3 parece describir un halo alrededor del Sol, separado de él por una distancia angular de 15 grados.

Los dos espectrómetros instalados en el vehículo tienen por misión determinar la distribución de masas de los núcleos de rayos cósmicos, "pesándolos" de uno en uno con una precisión superior al uno por ciento. El principio de operación de ambos instrumentos lleva consigo la medida de la energía cinética, E, y de la velocidad v de un núcleo incidente, así como la deducción de la masa m a partir de la ecuación de la energía cinética: $E = 1/2\, m\, v^2$. Por tanto, m es igual a $2E/v^2$, donde la masa se expresa en unidades de nucleones, es decir, de masas de un protón o un neutrón. Se lleva la cuenta de los núcleos que se mueven con velocidades entre un 25 y un 75 por ciento de la lumínica. Para calcular la masa nuclear con una exactitud del uno por ciento, ha de medirse la velocidad del núcleo con una precisión de décimas por ciento.



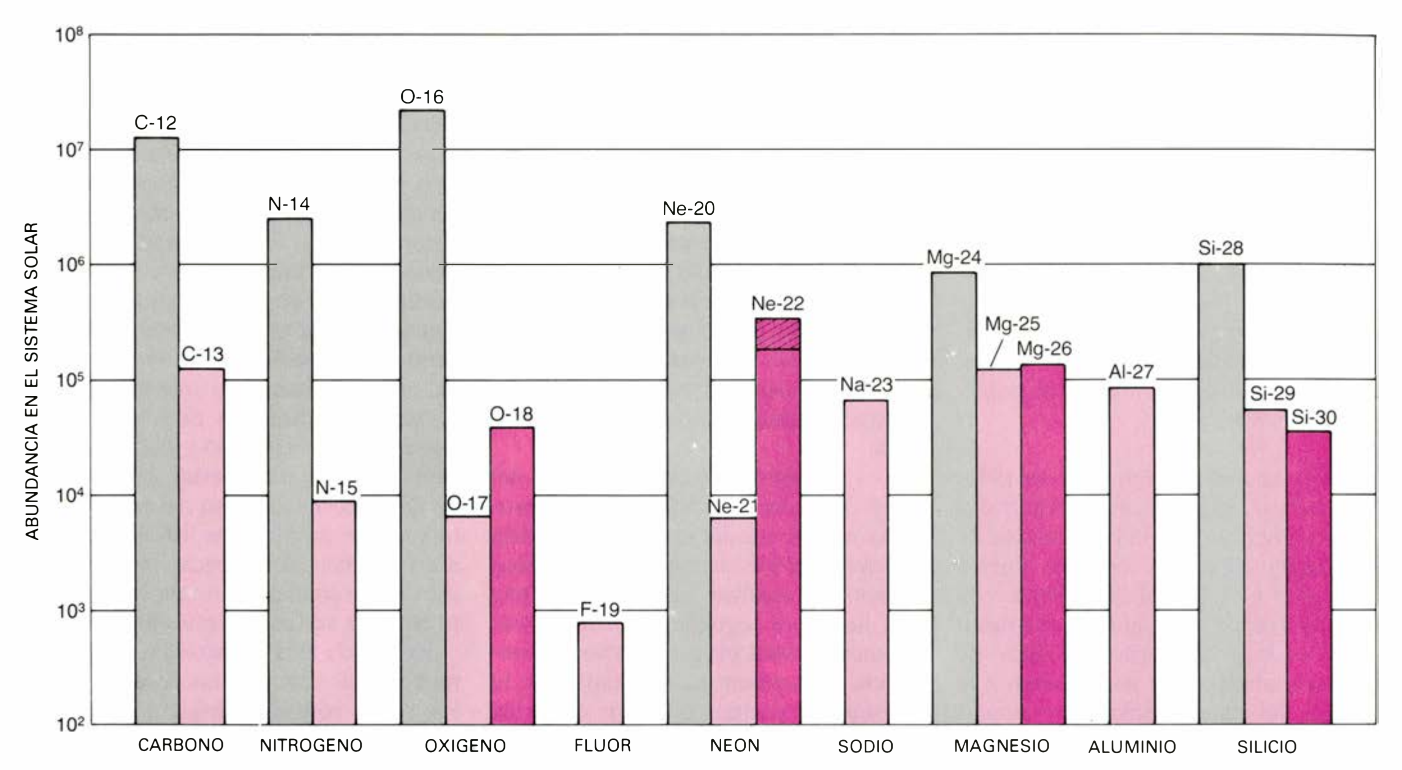
1. ESTE RESTO DE SUPERNOVA en la constelación Casiopea, Cas A, constituye un ejemplo de una región de nuestra galaxia donde las partículas nucleares se aceleran hasta grandes velocidades para convertirse en rayos cósmicos. En esta imagen de rayos X coloreada según una clave (obtenida con un detector montado en el satélite denominado Observatorio Einstein), el anillo brillante de intensa emisión de rayos X (*rojo y amarillo*) está formado por una capa en expansión de material estelar expulsada por la explosión de supernova, mezclado con gas interestelar. La capa mayor (*azul*) es, presumiblemente, gas interestelar calentado por el frente de la onda de choque de la explosión, que se mueve hacia fuera. Se cree que tales ondas de choque aceleran los rayos cósmicos. Basándose en la distancia estimada de Cas A, de 9000 años-luz, se supone que el diámetro del anillo brillante mide unos nueve años-luz. Del tamaño y la velocidad de expansión del anillo se infiere que la señal de explosión debió habernos llegado hace unos 325 años, pero no hay constancia de que se hubiera observado en la Tierra. Es el más joven resto de supernova conocido en nuestra galaxia. En las explosiones de supernova, los elementos pesados formados en estrellas de gran masa se dispersan por el gas interestelar, cambiando la composición de la galaxia a lo largo del tiempo. La imagen de rayos X fue obtenida por Stephen S. Murray y sus colaboradores en el Centro de Astrofísica del Harvard College Observatory y del Smithsoniano.



2. TODOS LOS ISOTOPOS NATURALES, o especies nucleares, de los nueve elementos desde el carbono hasta el silicio están representados esquemáticamente en este dibujo. Tres de los elementos (flúor, sodio y aluminio) tienen un solo isótopo estable. Los seis de la primera columna, dotados de igual número de protones (*negro*) y neutrones (*circunferencias*), son los que se sintetizan más fácilmente en las estrellas; se trata, por tanto, de los miembros de sus especies que se dan en mayor cuantía. Los isótopos de la segunda columna tienen un neutrón extra (*círculo en color*) y los de la tercera columna tienen dos neutrones extra. Por resultar de síntesis más difícil, los isótopos con neutrones extra son menos abundantes. A todos los núcleos de la primera columna, salvo el nitrógeno 14 (N-14), se les denomina también núcleos de partículas alfa, porque se pueden considerar formados por partículas alfa (núcleos de helio 4, compuestos por dos neutrones y dos protones). Los núcleos llevan una carga positiva *Z* igual al número de protones. Los protones y los neutrones reciben el nombre colectivo de nucleones. (Ilustración de Gabor Kiss.)

Los elementos sensores básicos de ambos instrumentos son detectores de estado sólido: discos de silicio, muy finos, de unos 10 centímetros cuadrados de superficie. Cuando un núcleo de gran energía atraviesa los discos, la carga eléctrica positiva del núcleo libera algunos electrones de la red cristalina del silicio. Como el núcleo pierde energía, que cede a los electrones, su velocidad disminuye. Aplicando una diferencia de potencial entre electrodos depositados en las caras del detector, se pueden recoger los electrones liberados, lo que despierta una débil señal eléctrica proporcional a la energía cinética perdida por el núcleo. Cuantos más protones haya en el núcleo, mayor será su carga positiva y mayor energía cederá a los electrones. Por otra parte, cuanto mayor sea la velocidad del núcleo, tanto menor energía cederá, porque tardará menos tiempo en cruzar el disco de silicio. La pérdida total de energía del núcleo al atravesar un detector viene a ser proporcional a Z^2/v^2, es decir, al cuadrado de la carga dividido por el cuadrado de la velocidad. Conforme el núcleo pase a través de sucesivos detectores, continuará decelerándose y perdiendo energía hasta detenerse. La suma de las señales eléctricas procedentes de la pila de detectores es una medida de la energía cinética que inicialmente llevaba el núcleo.

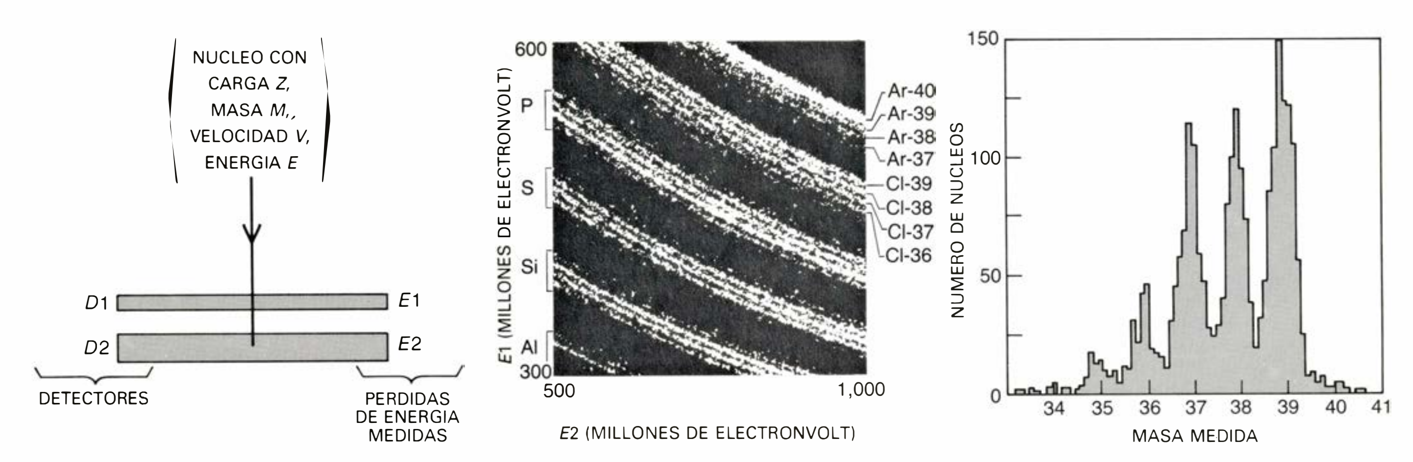
En principio, los valores de Z, m y E se pueden determinar únicamente para núcleos individuales de rayos cósmicos que llegan a detenerse en el espectrómetro. La energía perdida en el espectrómetro también depende, sin embargo, de la longitud del camino del núcleo en el instrumento, que varía según la dirección de donde procede el núcleo. Los núcleos de rayos cósmicos llegan uniformemente de todas las direcciones. Si se desea medir la masa de un núcleo con error inferior al uno por ciento, se deberá determinar la longitud de su recorrido con una precisión superior a esa cifra. La razón de por qué los instrumentos de Berkeley y del Cal Tech están capacitados para identificar isótopos de elementos, en tanto que los instrumentos anteriores de diseño similar sólo identificaban elementos, hay que buscarla en los detectores sensibles a la dirección incorporados en la ventana de entrada del instrumento. En el instrumento del Cal Tech, las trayectorias se determinan mediante una combinación de dos detectores de estado sólido sensibles a la dirección; el de Berkeley confía para ello en seis cámaras de "deriva" llenas de gas.



3. ABUNDANCIAS RELATIVAS en el sistema solar de los elementos desde el carbono hasta el silicio. Reflejan la distribución de elementos en el gas interestelar a partir del cual se condensaron hace cinco mil millones de años el Sol, los planetas y los demás cuerpos del sistema solar. Los isótopos se han inscrito en una escala logarítmica, donde la abundancia del silicio 28 (Si-28) se ha tomado como 10^6. Las abundancias se basan en medidas de la composición de cuerpos sólidos en el sistema solar, más que en la del Sol. La compilación fue realizada por A. G. W. Cameron. Los valores sirven de patrones, respecto a los cuales se puede medir la distribución de isótopos en los rayos cósmicos. Cualquier diferencia entre la materia del sistema solar y la materia de los rayos cósmicos cabe atribuirla a diferencias en las condiciones estelares bajo las cuales se sintetizaron las muestras. La abundancia de neón 22 en el sistema solar es incierta (*área rayada*), porque tanto los meteoritos como las partículas emitidas por el Sol muestran varias composiciones isotópicas diferentes.

Como el ISEE-3 gira alrededor del Sol y no de la Tierra y permanece siempre fuera del campo magnético terrestre, puede observar continuamente rayos cósmicos de gran energía procedentes de la galaxia. Cuando un rayo cósmico se detiene en uno de los espectrómetros, el instrumento registra las señales procedentes de los distintos detectores para su ulterior transmisión a la Tierra. Los espectrómetros responden a un amplio margen de elementos, desde el hidrógeno hasta el níquel, que tiene 28 protones. Colaboran con nosotros en estos experimentos Douglas E. Greiner y Harry H. Heckman, del Laboratorio Lawrence de Berkeley, y John D. Spalding y Rochus E. Vogt, del Cal Tech. En lo que sigue nos ceñiremos a nuestras medidas de isótopos del neón, magnesio y silicio en razón de su especial importancia para el estudio de la nucleosíntesis en las estrellas.

El neón, el magnesio y el silicio, res-



4. EL ESPECTROMETRO DE MASAS proyectado para análisis de rayos cósmicos identifica los elementos y los isótopos midiendo la pérdida de energía de núcleos individuales cuando son frenados y detenidos por dos detectores de estado sólido (*izquierda*). El instrumento produce señales eléctricas proporcionales a la energía perdida en cada detector. A grandes rasgos, el espectrómetro identifica los elementos porque su pérdida de energía viene a ser proporcional a Z^2/v^2, es decir, al cuadrado de la carga eléctrica dividido por el cuadrado de la velocidad. Los isótopos del mismo elemento pueden separarse porque su energía cinética ($E = 1/2\ m\ v^2$) es proporcional a la masa. La fotografía central de la ilustración recoge la separación de isótopos que se consigue con el instrumento construido en el Instituto de Tecnología de California durante una calibración llevada a cabo en el Laboratorio Lawrence de Berkeley. La fuente era un haz de argon 40 (Ar-40) suministrado por el acelerador Bevalac. Muchos de los núcleos del argon 40 se fragmentaron en núcleos más ligeros al chocar con el blanco. El conjunto de núcleos atravesaba un detector de 0,5 milímetros de espesor y lo detenía un segundo detector, de 1,7 milímetros de espesor. Cada núcleo, de los más de 10.000 que existen desde el aluminio ($Z = 13$) hasta el argon ($Z = 18$), aparece como un punto en una de las seis bandas anchas. Las franjas estrechas dentro de cada banda corresponden a isótopos separados de cada elemento. Las medidas habrán de tener una precisión diez veces mayor si se quieren identificar isótopos, no ya elementos. A la derecha, la resolución de los cinco isótopos del cloro.

pectivamente con 10, 12 y 14 protones, tienen cada uno tres isótopos estables. En la materia del sistema solar, los isótopos dominantes son el neón 20, magnesio 24 y silicio 28, núcleos compuestos de igual número de protones que de neutrones. Pueden por tanto considerarse compuestos respectivamente por cinco, seis y siete partículas alfa, o núcleos de helio (He-4) formados por dos protones y dos neutrones. Tales "núcleos de partículas alfa", muy estables, se fabrican copiosamente en los procesos estelares normales.

Hacia mediados del decenio de 1970, ciertos experimentos realizados por investigadores de la Universidad de New Hampshire, el Centro de Vuelos Espaciales Goddard de la NASA y la Universidad de Chicago demostraron que el neón, el magnesio y el silicio de los rayos cósmicos se asemejaban a la materia del sistema solar, por cuanto en ellos dominaban los mismos isótopos de partículas alfa. No era ningún descubrimiento inesperado, en efecto; pero sí importante: demostró que la mayoría de los rayos cósmicos no se producían en virtud de procesos anómalos de nucleosíntesis. Sin embargo, para extraer información más cuantitativa acerca de la naturaleza diferente o no de la nucleosíntesis de los rayos cósmicos y la propia de la materia solar, es preciso medir las abundancias relativas de los isótopos más pesados y más raros del neón, magnesio y silicio, es decir, los que contienen uno o dos neutrones extra.

Los espectrómetros a bordo del ISEE-3 fueron los primeros que proporcionaron medidas con resolución suficiente de esos isótopos ricos en neutrones. Al analizar los resultados hay que hacer en seguida una corrección: algunos núcleos observados son secundarios, o fragmentos, resultantes de la escisión de núcleos de rayos cósmicos más pesados que han chocado con átomos del gas interestelar. El porcentaje de secundarios en la muestra total se calcula observando las abundancias de ciertos núcleos que son relativamente raros entre los productos de nucleosíntesis y, por tanto, deberían estar ausentes a menos que se hayan formado por fragmentación. Entre esos núcleos citaremos los de litio, berilio y boro (respectivamente con tres, cuatro y cinco protones) e isótopos tales como el oxígeno 17 y el neón 21. Con la excepción del neón 21, parece ser que tres cuartas partes de los isótopos ricos en neutrones del neón, magnesio y silicio son núcleos vírgenes, o primarios, de la fuente de rayos cósmicos; han sobrevivido intactos tras un viaje de 10 millones de años a través del espacio interestelar; sólo aproximadamente una cuarta parte de ellos se forman por choques.

Realizada esta corrección, encontramos que las abundancias de isótopos en los rayos cósmicos difieren en forma importante de los que se encuentran en la materia del sistema solar. Se distinguen, sobre todo, por el predominio del neón 22 en los rayos cósmicos. En las muestras de materia del sistema solar hay de siete a doce núcleos de neón 22 por cada 100 núcleos de neón 20; en la fuente de rayos cósmicos, unos 50 núcleos de neón 22 por cada 100 núcleos de neón 20. Además, los cuatro isótopos ricos en neutrones del magnesio y silicio (Mg-25, Mg-26, Si-29 y Si-30) abundan en la fuente de rayos cósmicos en una proporción del 60 por ciento superior a la que se descubre en muestras de materia del sistema solar. Aunque del exceso de neón 22 en los rayos cósmicos habían informado ya los investigadores de New Hampshire, Goddard y Chicago, los primitivos experimentos carecían de sensibilidad suficiente para detectar los incrementos, menores, de los isótopos ricos en neutrones del magnesio y silicio.

El tamaño de las anomalías en los rayos cósmicos, que va desde el 60 por ciento a varios centenares por ciento, resulta impresionante si se tiene en cuenta que las anomalías de isótopos halladas en años recientes en algunos meteoritos se acercan al 1 por ciento, si llegan, en la relación de abundancias de dos isótopos, por ejemplo magnesio 25 y magnesio 24. En otras muestras de materia del sistema solar, rompe esa virtual uniformidad total la relación del neón 22 al neón 20, que varía en casi un factor de dos. La interpretación más razonable para las grandes anomalías que se registran en los rayos cósmicos sostiene que muchos núcleos que acaban como rayos cósmicos se sintetizaron

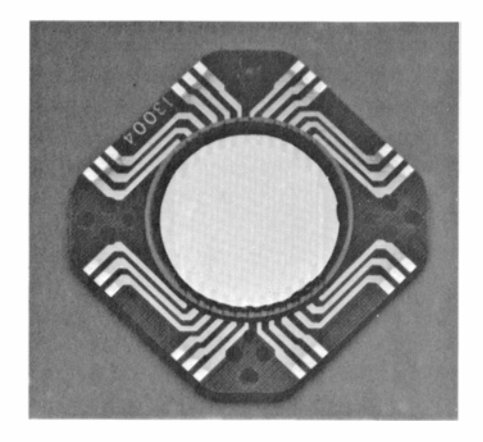
5. DETERMINACION DE LAS TRAYECTORIAS de los rayos cósmicos, factor esencial para el funcionamiento de un espectrómetro de isótopos en el espacio, donde los rayos cósmicos llegan de todas las direcciones. La energía que el núcleo incidente pierde al atravesar la pila de detectores depende de la longitud del camino recorrido por la partícula en el instrumento. La información suministrada cuando un núcleo atraviesa dos detectores sensibles a la dirección (P_1, P_2), permite calcular la trayectoria.

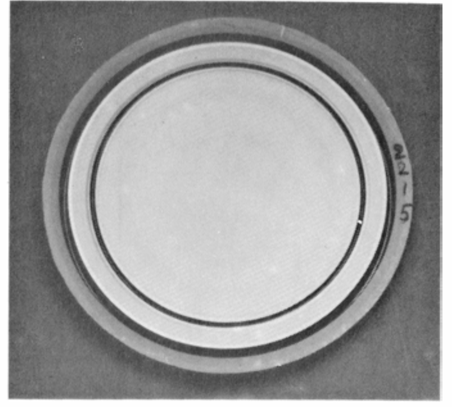
bajo condiciones distintas de las que crearon los núcleos típicos del sistema solar. Para comprender cómo pudieron surgir tales diferencias, necesitamos considerar algunos de los procesos que ocurren en la nucleosíntesis estelar.

Los elementos se sintetizan en las estrellas cuando núcleos más ligeros se funden entre sí para dar otros más pesados. El proceso de fusión convierte masa en energía cuando los núcleos más ligeros se reúnen para formar elementos tan pesados como el hierro; de ahí que, al describirlo, se hable también de "combustión" de los núcleos más ligeros. Ese proceso de fusión comprende dos fases distintas. El grueso de los núcleos pesados se forma mediante la lenta y gradual fusión de núcleos progresivamente más pesados en las estrellas ordinarias, proceso que libera energía termonuclear a lo largo de millones o miles de millones de años.

En las estrellas de masa superior a cierto límite que han agotado su provisión de núcleos capaces de liberar energía termonuclear mediante la fusión actúa una clase de nucleosíntesis claramente diferente. Desprovistas de una fuente persistente de energía, las estrellas se contraen y después explotan constituyendo una supernova. En esta fase final explosiva, alimentada por la liberación de energía gravitatoria, se sintetiza, literalmente como un rayo, un amplio espectro de núcleos pesados. El grueso de la masa de la estrella sale expulsado y dispersado hacia el medio interestelar circundante, dejando detrás un denso y compacto objeto, una estrella de neutrones o un agujero negro, verosímilmente.

Como las fases gradual y explosiva de la nucleosíntesis se caracterizan por escalas de temperatura y tiempo ampliamente distintas, originan núcleos muy diferentes. La mayoría de los núcleos con igual número de protones y neutrones se sintetizan en una sucesión de procesos de combustión lenta, empezando por la fusión de núcleos de hidrógeno (protones) para formar núcleos de helio ordinario, He-4. En tales procesos, la mitad de los protones se convierten en neutrones por desintegraciones radiactivas en las que la carga positiva de un neutrón se emite en forma de electrón positivo, o positrón. Los núcleos de hidrógeno también se combinan con cualquier traza de núcleos de carbono, nitrógeno u oxígeno (a los que colectivamente nos referiremos como núcleos CNO) que pudiera haber existido ya en la materia de la que nació la estrella. Para que hubiera





6. EL DETECTOR SENSIBLE A LA POSICION en la fotografía de la izquierda tiene 24 láminas metálicas paralelas sobre una superficie de la oblea de silicio y, perpendicularmente a ellas, otro conjunto de 24 láminas en la otra superficie. El detector suministra así un conjunto de coordenadas *x-y* para un rayo cósmico incidente. Mediante dos de tales detectores, uno encima del otro, el espectrómetro del Instituto de Tecnología de California determina las trayectorias de los núcleos incidentes. La fotografía de la derecha nos muestra uno de los 10 detectores para medir la energía que están apilados bajo los detectores sensibles a la posición. El detector tiene dos zonas de detección (*gris claro*). El anillo que rodea al disco central sirve para detectar núcleos que entran o salen a través de la cara lateral del instrumento. El espectrómetro del Cal Tech es uno de los dos que transporta el Tercer Explorador Internacional Sol-Tierra (ISEE-3). El segundo instrumento, construido en la Universidad de California en Berkeley, incorpora detectores medidores de energía similares, aunque con un esquema diferente para determinar las trayectorias seguidas.

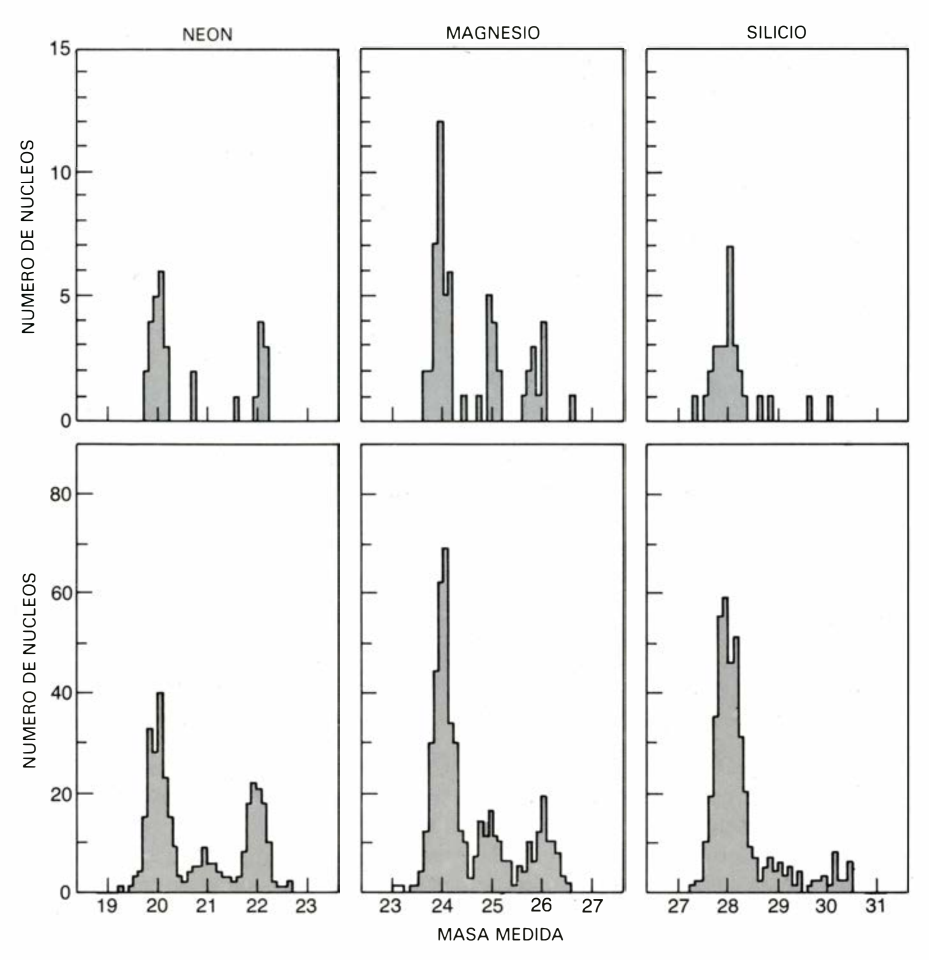
núcleos de CNO presentes, debieron crearse en una generación anterior de nucleosíntesis estelar y ser inyectados en el medio interestelar que originó la estrella. En la fase de combustión de hidrógeno, los núcleos de carbono y oxígeno se convierten en nitrógeno ordinario, N-14, a través de la serie de reacciones nucleares conocidas por ciclo CNO.

Nuestro Sol transforma hidrógeno en helio y continuará haciéndolo otros cinco mil millones de años más. Una estrella de masa mayor quemará su hidrógeno a una velocidad superior porque su temperatura y densidad internas son mayores. Al final de la fase de combustión de hidrógeno, el núcleo de una estrella de gran masa se contrae; la temperatura se eleva hasta quemar helio y producir principalmente carbono 12 y oxígeno 16 mediante la fusión, respectivamente, de tres y cuatro núcleos de helio. En esta fase de la vida de la estrella, el helio también se combina con el nitrógeno 14 en una serie de reacciones de fusión y subsiguientes desintegraciones radiactivas cuyo producto final es el neón 22, núcleo que tiene dos neutrones más que protones.

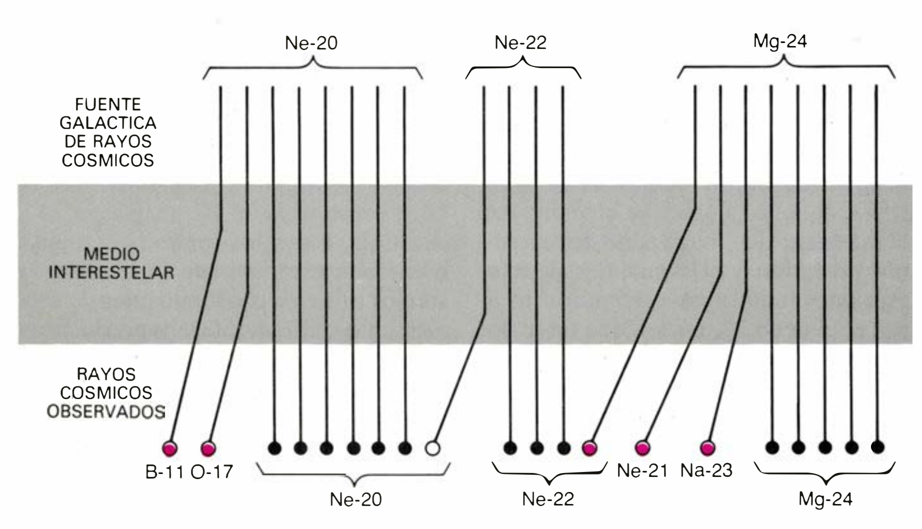
El número total de neutrones extra en una estrella importa mucho, ya que rige la producción de otros núcleos ricos en neutrones sintetizados tanto en la fase gradual como en la fase final explosiva de la vida de la estrella. El neón 22 se obtiene del nitrógeno 14, que a su vez se recaba de la provisión original de núcleos CNO en la estrella, de modo que la abundancia de neón 22 es proporcional a la cantidad original de CNO. El 2 por ciento de la materia del sistema solar está constituida por núcleos CNO. Por consiguiente, una estrella que originariamente estaba provista de la misma abundancia de núcleos CNO acabará con un 2 por ciento de neón 22 en su núcleo tras completar el proceso de combustión de hidrógeno y helio. Aunque cada núcleo de neón 22 tiene un exceso de neutrones de casi el 10 por ciento, el grueso de la materia estará constituido por carbono 12 y oxígeno 16, núcleos que carecen de neutrones extra, de modo que el exceso global de neutrones será sólo del 0,2 por ciento, correspondiendo a dos neutrones más que protones por cada 1000 nucleones.

Cuando el helio de la estrella está próximo a agotarse, ésta se contrae de nuevo y se hace más caliente, posibilitando la fusión de núcleos de carbono y oxígeno para formar núcleos tales como el neón 20, magnesio 24 y silicio 28, todos los cuales tienen igual número de protones que de neutrones y son los isótopos predominantes de esos elementos. Dichos núcleos productores de partículas alfa son, a su vez, los principales productos de la fase gradual de la nucleosíntesis. Algo del neón 22 se convierte también en magnesio 25 en esta fase, liberando neutrones que pueden ser capturados por otros núcleos para formar isótopos adicionales ricos en neutrones.

A medida que el núcleo de una estrella de gran masa va agotando la provisión de núcleos ligeros capaces de alimentar la combustión gradual, la



7. **DISTRIBUCION DE ISOTOPOS de rayos cósmicos según la determinaron, en el espacio, el espectrómetro del Cal Tech (*arriba*) y el de Berkeley (*abajo*). Nos enseña que el neón, magnesio y silicio están representados cada uno por sus tres isótopos estables. Como en el sistema solar, predominan los isótopos constituidos por partículas alfa (neón 20, magnesio 24 y silicio 28). Los 120 sucesos de la muestra del Cal Tech se recogieron en los primeros tres meses y medio de observaciones, antes de que el espectrómetro quedara parcialmente inutilizado. Los 1226 sucesos de la muestra de Berkeley representan los datos acumulados a lo largo de un período de dos años, después de los cuales se averiaron los sensores de trayectorias. Por consiguiente, al combinar los resultados se da mayor peso a los datos de Berkeley. Los experimentos fueron llevados a cabo por los autores y sus colaboradores, entre ellos Douglas E. Greiner y Harry H. Heckman, del Lawrence de Berkeley, y John D. Spalding y Rochus E. Vogt, del Cal Tech.**



8. **FRAGMENTACION DE RAYOS COSMICOS por choques en el medio interestelar. Dicha partición altera la composición de los núcleos originalmente acelerados en la fuente de rayos cósmicos. A los fragmentos que resultan de las colisiones se les califica de "secundarios". Se recogen algunos de los secundarios (*en color*) que pueden resultar cuando núcleos primarios de tres isótopos (neón 20, neón 22 y magnesio 24) chocan con átomos de gas interestelar. Isótopos tales como el boro 11, el oxígeno 17 y el neón 21 escasean entre los productos de nucleosíntesis estelar; pueden considerarse secundarios. Salvo el neón 21, tres cuartas partes de los núcleos ricos en neutrones de neón, magnesio y silicio son de origen primario.**

estrella se hace inestable y acaba por producirse una explosión de supernova, lanzando una onda de choque desde el interior profundo de la estrella y a través de la materia que la recubre y calentándola a varios miles de millones de grados durante décimas de segundo. En este breve período se sintetizan muchos núcleos adicionales que comprenden el magnesio 25, magnesio 26, silicio 29 y silicio 30. Las cantidades de esos isótopos ricos en neutrones están determinadas por la abundancia de neón 22, que es una fuente de neutrones extra, antes de la explosión. Como resultado, el exceso de neutrones, originalmente determinado por la abundancia de CNO en la estrella, rige la proporción de núcleos ricos en neutrones que se sintetizan tanto en el período gradual como en el explosivo de la vida de la estrella.

Hace algunos años, W. David Arnett, quien actualmente se halla en la Universidad de Chicago, sugirió que la galaxia debería mostrar un continuo aumento del número de núcleos ricos en neutrones en virtud del efecto acumulativo de la nucleosíntesis en las sucesivas generaciones de estrellas. Los rayos cósmicos que llegan a la vecindad de la Tierra hoy son presumiblemente mucho más jóvenes que la materia que forma el sistema solar, a la que se le supone cinco mil millones de años de edad; sería de esperar, pues, que los rayos cósmicos reflejasen el aumento de núcleos ricos en neutrones. Pero los modelos de evolución galáctica en vigor sugieren que el esperado enriquecimiento en isótopos ricos en neutrones de la galaxia entera, a lo largo de los últimos 5000 millones de años, quizá no sea suficiente para explicar nuestras observaciones de rayos cósmicos, particularmente en el caso del neón 22.

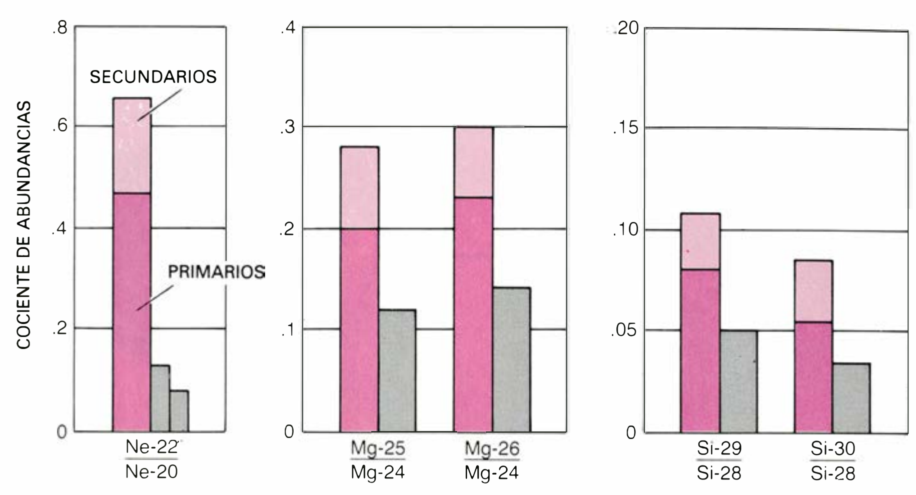
Más recientemente, S. E. Woosley, de la Universidad de California en Santa Cruz, y T. A. Weaver, del Laboratorio Nacional Lawrence Livermore, han propuesto que los excesos observados de isótopos ricos en neutrones en los rayos cósmicos galácticos podrían explicarse con la condición siguiente: que en nuestra galaxia hubiera regiones donde la abundancia de CNO fuese significativamente mayor que la de la materia del sistema solar. De sus cálculos se infiere que la materia expulsada por las supernovas, cuya abundancia inicial de CNO fuese un 80 por ciento mayor que en el sistema solar, se enriquecería en un 80 por ciento en neón 22 y de 50 a 70 por ciento en las abundancias de cada uno de los isótopos de magnesio y silicio ricos en neutrones. Confirma es-

te modelo el casi igual enriquecimiento que encontramos para cuatro de los cinco isótopos ricos en neutrones. Por tanto, los excesos de 60 por ciento de magnesio 25, magnesio 26, silicio 29 y silicio 30 reciben una explicación adecuada si los suponemos resultantes de un 80 por ciento de aumento en la abundancia de CNO en las regiones de donde surgen los rayos cósmicos.

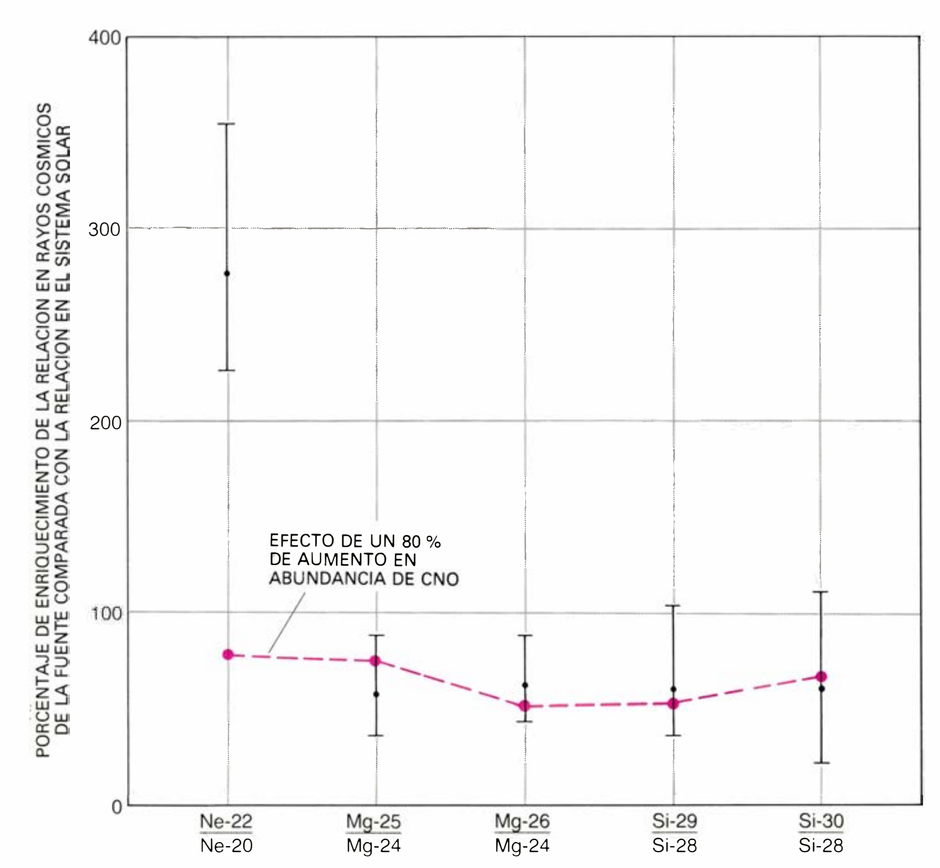
Sin embargo, el exceso de neón 22 que encontramos es superior al 200 por ciento. Contrasta netamente con el 80 por ciento de enriquecimiento que cabría esperar de un 80 por ciento de aumento en la abundancia galáctica de CNO. El exceso observado de neón 22 puede obedecer a algún aspecto de la nucleosíntesis que no se ha tenido en cuenta en los sencillos modelos de la evolución química de la galaxia, o quizá se deba también a un conocimiento impreciso de las abundancias de isótopos del neón en la materia del sistema solar. La verdad es que se han medido composiciones isotópicas diferentes en meteoritos y en el Sol.

Otra explicación plausible del gran exceso de neón 22 en los rayos cósmicos es que éstos no procedan de una muestra representativa de la materia interestelar, sino de una clase limitada de objetos estelares que sinteticen neón 22 con una abundancia muy superior a la encontrada en el sistema solar. Cierto número de investigadores han propuesto modelos en los que se podría sintetizar un gran exceso de neón 22, al tiempo que se mantendría una mezcla de los otros núcleos casi igual a la del sistema solar. No es todavía evidente, sin embargo, que tales fuentes justificaran también las demás anomalías isotópicas que hemos descubierto en el magnesio y silicio.

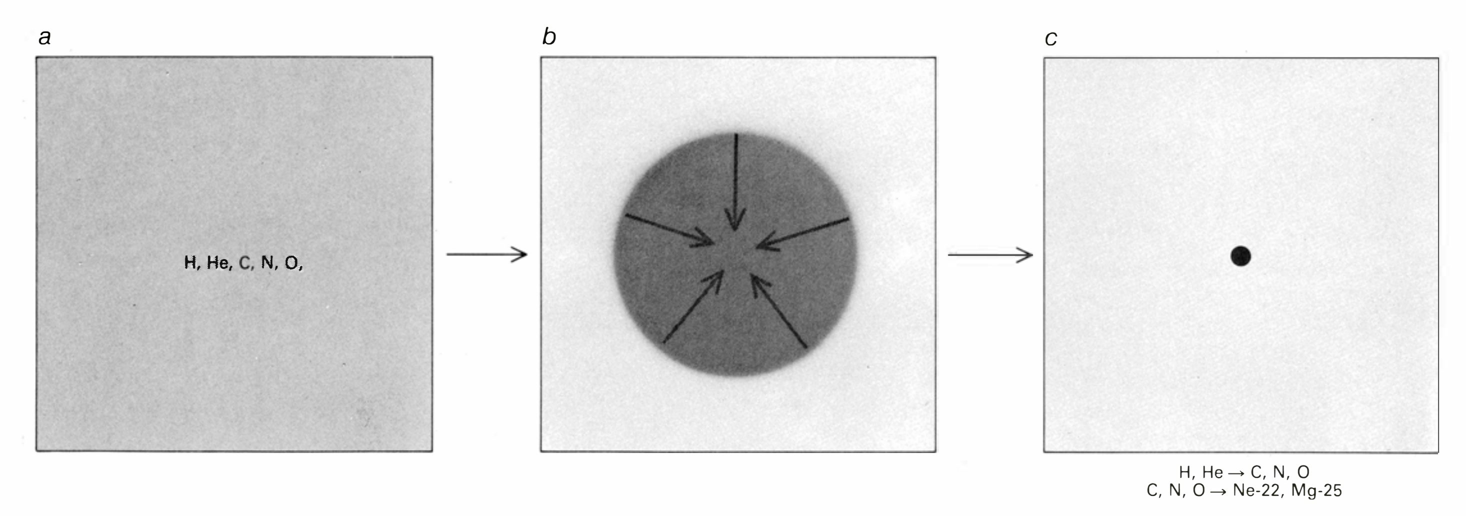
K. A. Olive y David N. Schramm, de la Universidad de Chicago, acaban de proponer una hipótesis más radical. En su opinión, los rayos cósmicos reflejan con minuciosa exactitud la composición del medio interestelar de la galaxia en general, y que es el sistema solar el que es atípico. Aducen recientes estudios de meteoritos, según los cuales, cuando se formó el sistema solar, contenía un isótopo radiactivo de vida corta del aluminio, el Al-26. La existencia del aluminio 26 se infiere de la presencia en los meteoritos de su producto de desintegración, magnesio 26, en íntima asociación con el aluminio 27, que es estable. El aluminio 26 tiene una vida media de sólo un millón de años; difícilmente, pues, podría haber estado presente cuando el sistema solar



9. ABUNDANCIAS DE LOS ISOTOPOS RICOS EN NEUTRONES de los rayos cósmicos, representadas en color y en forma de cociente; los isótopos de partículas alfa más abundantes del neón, magnesio y silicio constituyen el denominador. Las abundancias correspondientes en materia del sistema solar se representan en forma de barras grises. En los cocientes para los rayos cósmicos se distingue entre los primarios y los secundarios. Como el neón del sistema solar muestra diversas relaciones de abundancia isotópica, el intervalo de variación se indica mediante dos barras. Las medidas de rayos cósmicos son la media ponderada de los resultados de la Universidad de California en Berkeley y el Instituto de Tecnología. Manifiestan que la fuente tiene un exceso de isótopos ricos en neutrones de estos elementos.



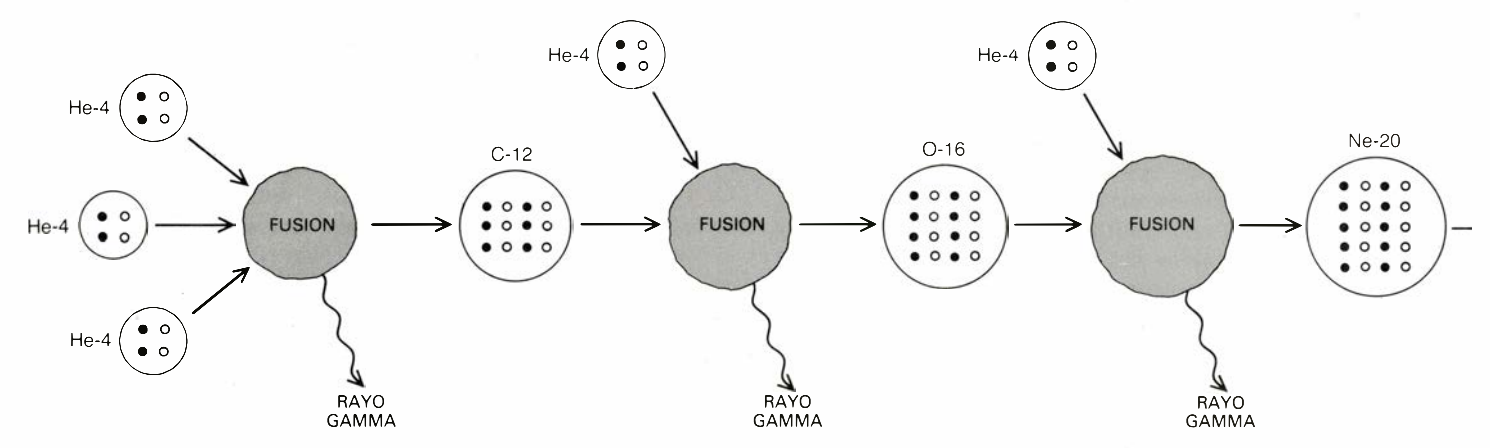
10. EXCESO DE ISOTOPOS RICOS EN NEUTRONES en la materia fuente de rayos cósmicos. Refleja presumiblemente los procesos de nucleosíntesis en estrellas cuya composición inicial difería de la composición de la generación de estrellas que proveyó la materia para el sistema solar. S. E. Woosley, de la Universidad de California en Santa Cruz, y T. A. Weaver, del Laboratorio Nacional Lawrence Livermore, han propuesto la hipótesis de que el exceso de isótopos ricos en neutrones se explicaría si las estrellas responsables de los núcleos de rayos cósmicos gozaran, en un comienzo, de una abundancia de núcleos de carbono, nitrógeno y oxígeno (denominados colectivamente núcleos CNO) un 80 por ciento mayor que las estrellas responsables de los núcleos del sistema solar. La línea de trazos en color señala el aumento de isótopos ricos en neutrones de neón, magnesio y silicio que debería observarse en los rayos cósmicos de tales estrellas. La línea de trazos pasa por cuatro de las cinco barras que representan el aumento observado en isótopos ricos en neutrones en los rayos cósmicos, comparado con esos isótopos en el sistema solar. El cociente del neón 22, rico en neutrones, al neón 20 discrepa, sin embargo, en más del 200 por ciento, aun cuando se tome como base de comparación el valor máximo del sistema solar. La longitud vertical de las distintas barras indica las incertidumbres experimentales en las medidas de rayos cósmicos.



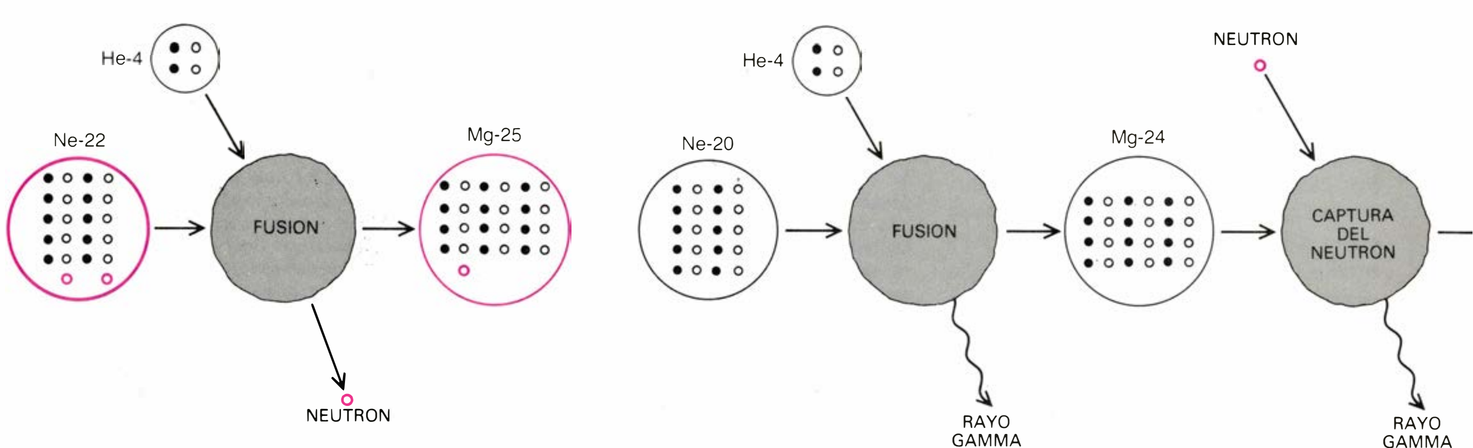
11. ABUNDANCIAS RELATIVAS de isótopos en la galaxia. Vienen determinadas por la formación, evolución y rotura explosiva de las estrellas de gran masa. En una región de la galaxia donde la composición del gas interestelar (*a*) sea similar a la de la materia del sistema solar, una nube de gas colapsa bajo la influencia de su propia gravedad, dando nacimiento a una nueva estrella (*b*). En el interior de ésta, la fusión termonuclear convierte parte del hidrógeno y helio originales en núcleos de partículas alfa: carbono 12 y oxígeno 16 (*c*). Simultáneamente, los núcleos de CNO que originariamente estaban presentes en el combustible estelar se convierten en núcleos ricos en neutrones más pesados: neón 22 y magnesio 25. Cuando esta combustión lenta ha agotado todo el combustible nuclear en el interior de la estrella, ésta explota como una supernova (*d*). La onda de choque generada por la explosión sintetiza

se condensó a partir de la materia interestelar, a menos que se hubiera producido recientemente. La explicación más probable es que una explosión de supernova inyectase el alumnio 26 inmediatamente antes de la formación del sistema solar. En ese caso, otros elementos, también inyectados por la supernova, podrían haber dado al sistema solar una composición apreciablemente diferente de la composición promedio del medio interestelar.

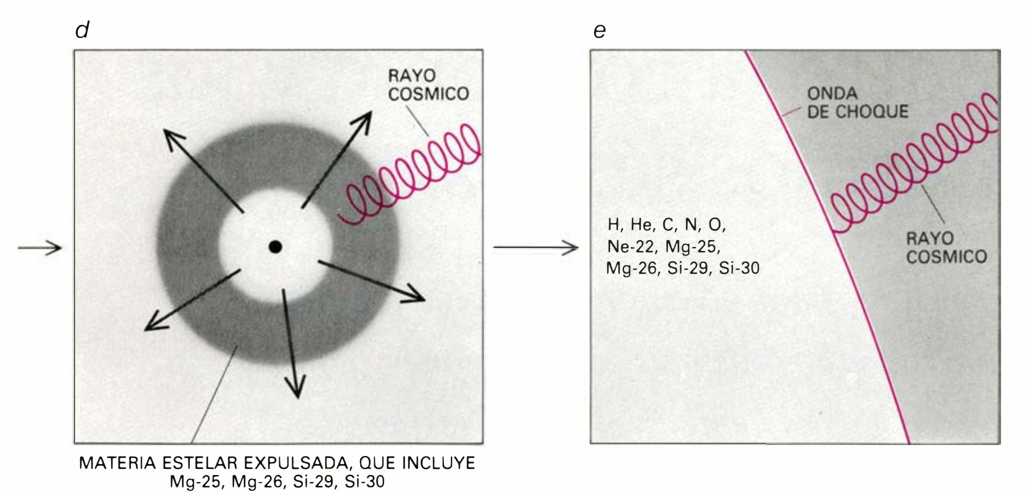
Olive y Schramm proponen un modelo específico en el que la concentración de núcleos ricos en neutrones típica del medio interestelar en el momento en que se formó el sistema solar quedó diluida por la inyección de un exceso de núcleos productores de partículas alfa procedentes de una fuente o fuentes de supernova. Estas abundancias modificadas discreparían, por consiguiente, de la abundancia medida en los rayos cósmicos, haciendo que éstos parecie-



12. FUSION LENTA de los núcleos de elementos ligeros para crear núcleos de masa mayor. En la primera fase (no representada), los núcleos de hidrógeno se funden para formar helio (partículas alfa) y esencialmente todos los núcleos de CNO que pudieran haber existido en el combustible original se convierten en nitrógeno 14. La secuencia de la izquierda ilustra la segunda fase, en la cual los núcleos de helio se funden para originar núcleos de partícu-



13. LA NUCLEOSINTESIS EXPLOSIVA se produce durante el estallido de supernova de una estrella de gran masa cuando ha terminado la fase de combustión lenta. La onda de choque explosiva calienta la materia estelar hasta varios miles de millones de grados C durante algunas décimas de segundo, iniciando un complejo conjunto de reacciones nucleares, de las que se ilustran algunas. Importa destacar la fusión del neón 22 y helio 4 (*izquierda*),

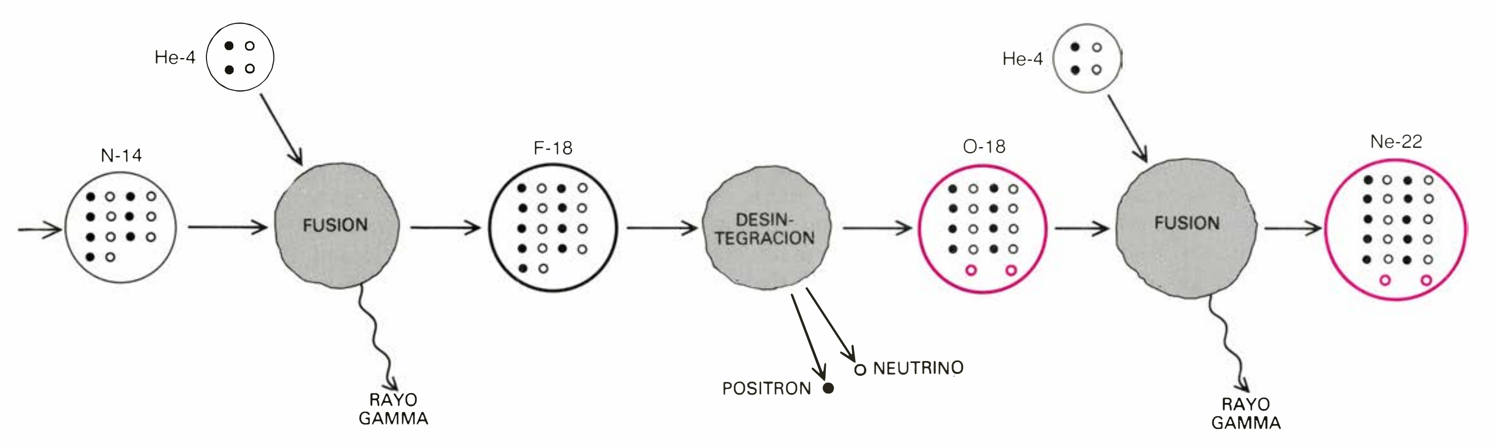


núcleos pesados adicionales y expulsa la mayoría de los productos de la nucleosíntesis hacia el gas interestelar. Las repeticiones de esta sucesión de acontecimientos en subsiguientes generaciones de estrellas enriquece sin cesar el gas interestelar en carbono, nitrógeno y oxígeno y en núcleos pesados con exceso de neutrones. Algunos núcleos se aceleran hasta alcanzar velocidades de rayos cósmicos, posiblemente por las ondas de choque de las supernovas (*e*). La aceleración de los rayos cósmicos podría producirse también de un modo directo cuando la supernova expulsa materia hacia el espacio interestelar (*d*).

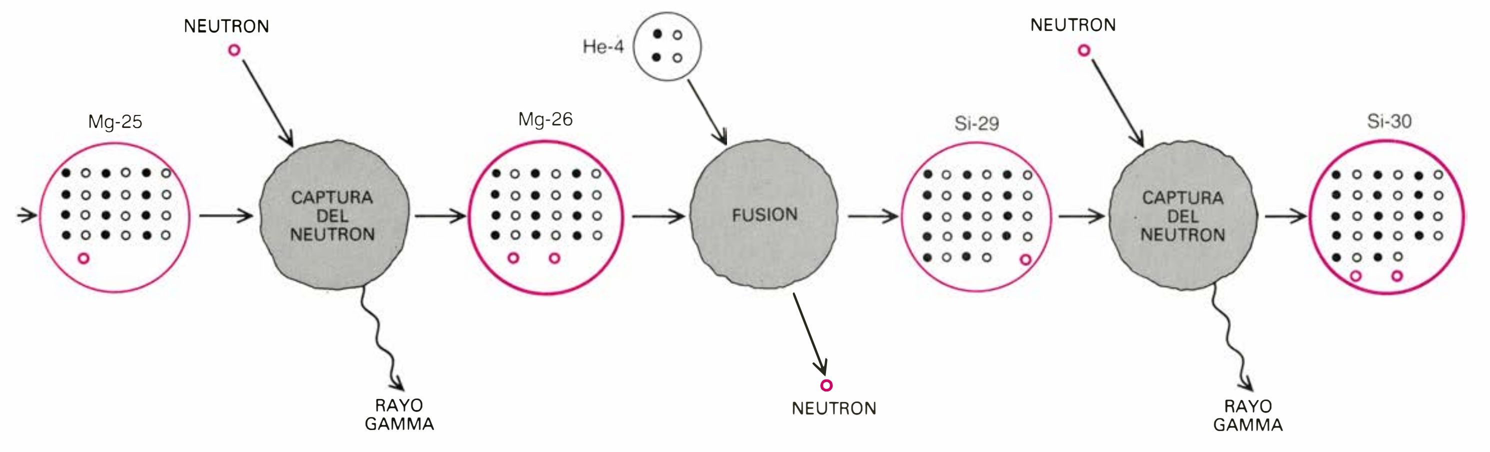
sen ricos en neutrones, en concordancia cualitativa con nuestras observaciones. Si Olive y Schramm tienen razón, la significación real de las observaciones de rayos cósmicos residiría en la luz que arrojen sobre los sucesos asociados a la formación del sistema solar.

Como revelan esos modelos y propuestas, la mezcla de los elementos e isótopos en una muestra de materia depende de la naturaleza de la nucleosíntesis en las estrellas y de la historia de la formación de estrellas y de su evolución en la galaxia entera. Nuestro conocimiento de tales procesos progresa y se extiende gracias a las observaciones detalladas de las muestras de materia que los rayos cósmicos traen de otras partes de la galaxia. Cuanto sabemos hasta ahora es todavía muy limitado. Sólo se han determinado con precisión las abundancias isotópicas de tres elementos –neón, magnesio y silicio– en la fuente de rayos cósmicos, y en cada caso se ha encontrado que difieren de las correspondientes abundancias en las muestras del sistema solar. Las anomalías observadas pueden muy bien ser la regla y no la excepción.

Otros elementos, así el hierro y el níquel, poseen varios isótopos estables cuyas abundancias servirían como sensibles sondas de la nucleosíntesis. Se sabe ya que el isótopo más abundante del hierro en el sistema solar y en los rayos cósmicos es el Fe-56. Todavía no hay medidas precisas, sin embargo, de las abundancias en los rayos cósmicos de otros dos isótopos del hierro, el Fe-54 y el Fe-58, que se sintetizan en diferentes proporciones bajo distintas condiciones estelares. Si los futuros vehículos espaciales llevan instrumentos capaces de medir estos y otros isótopos con precisión, ello podría ayudar a resolver los enigmas aquí planteados.



las alfa; así: carbono 12, oxígeno 16 y neón 20. La secuencia de la derecha muestra la fusión y los procesos de desintegración radiactiva que convierten los núcleos de nitrógeno 14 (que se han sintetizado antes a partir de núcleos de carbono) en núcleos de neón 22. La abundancia final de neón 22, que tiene dos neutrones más que protones, dependerá de la abundancia inicial de carbono, nitrógeno y oxígeno (CNO) que haya en el combustible de la estrella.



porque los neutrones liberados pueden ser capturados en la secuencia de reacciones (*derecha*) que dan isótopos ricos en neutrones de magnesio y silicio. Las abundancias de esos isótopos están, pues, determinadas por la provisión de neón 22, que depende de la abundancia inicial de NCO en la estrella. Aunque en la fase de combustión lenta tiene lugar una síntesis similar de isótopos ricos en neutrones, las velocidades de reacción son entonces mucho menores.

Interacción entre la corriente y los organismos sésiles

Multitud de organismos marinos viven fijados en los fondos litorales. Han desarrollado finísimas adaptaciones ante las fuerzas que ejercen la corriente y el batir de las olas

M. A. R. Koehl

Si paseamos por una costa marina rocosa durante la marea baja descubriremos animales y plantas, en número y variedad notables, que viven fijados al fondo soportando el batir del oleaje. Estos organismos sésiles, que corren el riesgo de que el agua en movimiento los arranque y arrastre, dependen, sin embargo, del agua para recabar sus propios nutrientes, librarse de sus excreciones y dispersar sus descendientes hacia nuevos puntos de anclaje. ¿Cómo se las arreglan?

Detengámonos y contemplemos atentamente el comportamiento de las plantas y los animales sésiles. (Vengo haciendo ese examen, desde hace años, por toda la costa pacífica de Norte y Sudamérica y los arrecifes de coral caribeños.) Los organismos encarnan varios compromisos entre la tendencia a maximizar y minimizar los efectos del flujo de agua. Los que viven en un medio de circulación lenta suelen presentar rasgos que compensan la falta de transporte por parte del agua. Crean, así, sus propias corrientes alimentarias o toleran niveles bajos de oxígeno. De forma similar, los organismos que viven inmersos en un flujo rápido poseen mecanismos que aumentan su capacidad de soportar y utilizar el agua que fluye precipitadamente a su alrededor. Las características que permiten a los organismos sésiles mantenerse en un determinado lugar son más variadas de lo que cabría presumir.

Varias preguntas se le plantean al investigador que desea explorar de qué modo un organismo sésil se adapta, desde el punto de vista mecánico, al agua en movimiento. ¿Cómo afecta la estructura del organismo al flujo de agua que pasa por su lado y a las fuerzas inducidas por el flujo que actúan sobre él? ¿En qué medida la forma del organismo determina la tensión que sufren sus tejidos cuando se halla sujeto a fuerzas de flujo? (Por tensión se entiende la fuerza o carga por unidad de superficie transversal del material que soporta la carga.) ¿De qué modo el material estructural de un organismo afecta a la manera en que éste se deforma o se rompe en respuesta a tensiones producidas por el flujo?

Los organismos marinos sésiles viven en varios tipos básicos de flujo. Existe una distinción fundamental entre hábitats intermareales y submareales, esto es, entre los hábitats situados en la frontera de la pleamar y la bajamar y los que quedan por debajo del límite de la bajamar. Una planta o animal intermareal, instalado en un punto donde las olas rompen sobre una costa expuesta, se halla sujeto a un flujo rápido y turbulento. He medido el flujo sobre organismos de este tipo y he encontrado que la velocidad alcanza su máximo durante la marejada que avanza hacia la costa y la resaca que se dirige hacia el mar después de romper una ola.

Para los organismos que viven sobre el fondo de la zona submareal superior de una costa batida por el oleaje, el flujo hacia uno y otro lado con el que se enfrentan cuando las olas pasan por encima de ellos es menor. (Recuérdese que la posición de un organismo con respecto a las olas que rompen, y con ello el tipo de flujo con el que se enfrenta, cambia a medida que la marea sube y baja.) Si un organismo vive a mayor profundidad que la mitad de la distancia entre las crestas de olas sucesivas que pasan por encima, no "nota" las olas. Este, así como los organismos que viven en bahías y canales protegidos, se halla en cambio expuesto a simples corrientes uniformes o a las corrientes uniformes que se invierten periódicamente de las mareas.

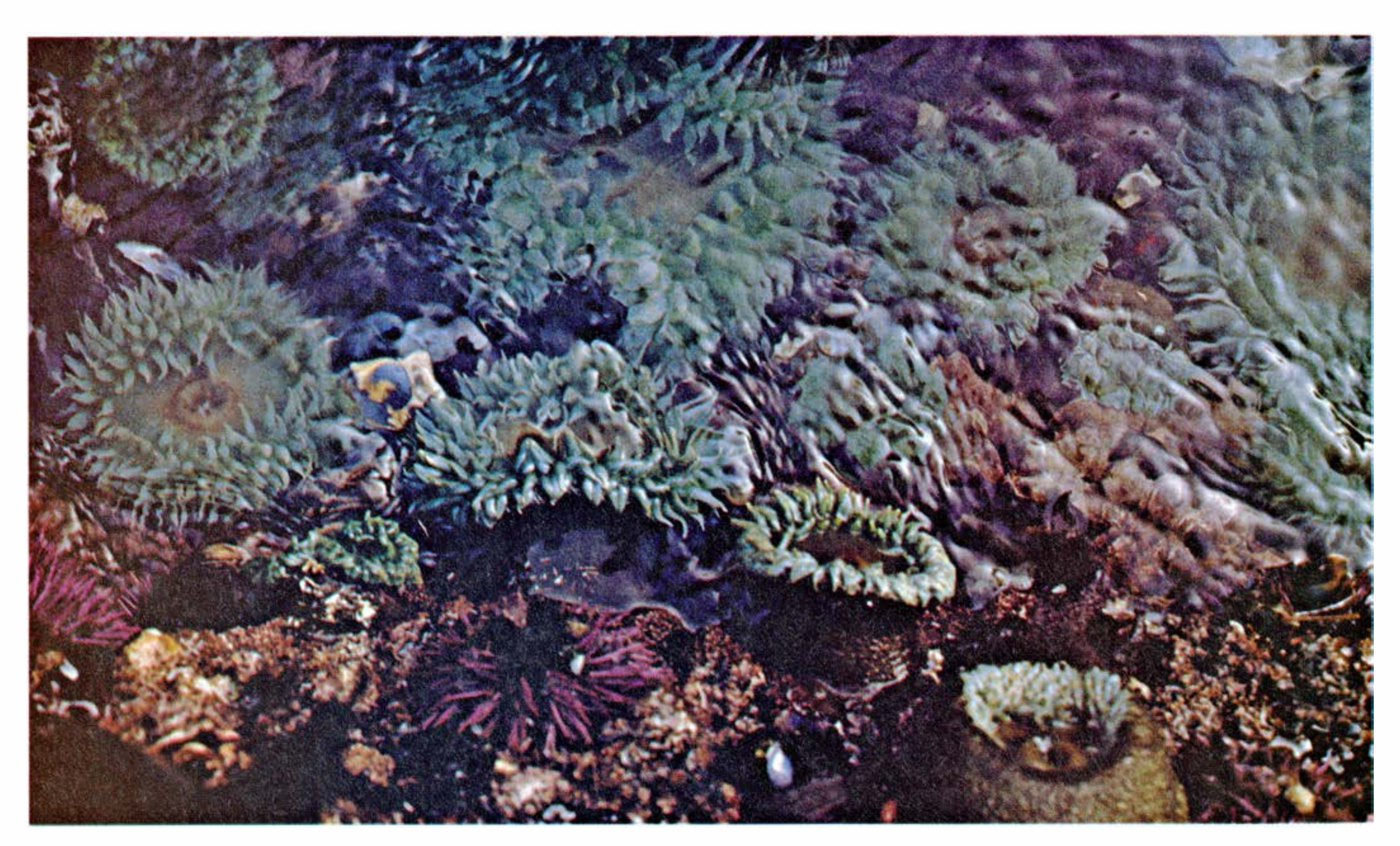
Cuanto he descrito hasta aquí podría llamarse macrohábitats de flujo. Cada macrohábitat contiene microhábitats locales donde corre un flujo distinto. A la hora de abordar un microhábitat debe tenerse en cuenta un rasgo característico de todo fluido que circule por una superficie sólida: se hace más lento en las inmediaciones de la superficie. Esa capa de fluido que avanza despacio se denomina capa límite; cuanto más lento sea el flujo o cuanto más alejado del borde de ataque de un objeto se detecte, mayor será la capa límite. En la capa límite viven organismos extremadamente pequeños (larvas que acaban de instalarse y plantas y animales unicelulares bentónicos, es decir, que viven en el fondo) y organismos incrustantes planos (algas crustáceas y algunos briozoos). El microhábitat de flujo de un organismo pequeño puede ser muy distinto del flujo principal que corre sobre la zona en que vive. Además, el ambiente marino no es por lo general plano y liso; las protuberancias y grietas, e incluso los demás organismos que se encuentran alrededor de una planta o animal grandes, pueden alterar notablemente el flujo que pasa sobre ellos.

Las anémonas de mar del género *Anthopleura* constituyen buenos ejemplos de animales que se desenvuelven en microhábitats de flujo sorprendentes. Proliferan en las costas rocosas, de California a la Columbia Británica. Las grandes anémonas verdes (*A. xanthogrammica*) tapizan el fondo de los canales de marejada intermareales en lugares particularmente expuestos. Adheridas al fondo de la zona intermareal inferior, viven ingiriendo mejillones y erizos de mar que son arrancados por las olas y arrastrados hasta su alcance.

He medido los flujos de la corriente principal de estos canales de marejada y los flujos inferiores, donde se encuen-



1. CINTURON ALGAL a lo largo de la costa rocosa de Chile central y austral, expuesto a un mar embravecido que impone fuertes tensiones en los tejidos de las plantas. La fuerza se aplica en un sentido cuando la ola rompe y en el contrario cuando el agua de la ola que acaba de romper vuelve al mar. Se trata en este caso del alga *Lessonia nigrescens*, que responde a las fuerzas doblándose con el flujo. Un alga acompañante en el cinturón, *Durvillea antarctica*, se dispone más o menos paralela al fondo del mar en respuesta al tirón de la corriente. La tensión máxima en el estipe, o pedúnculo, de un alga presionada por el flujo, como *D. antarctica*, es considerablemente menor que la del estipe de un alga que se dobla, lo que hace justamente *L. nigrescens*.



2. GRANDES ANEMONAS VERDES, *Anthopleura xanthogrammica*, tapizan el fondo de los canales de marejada en costas rocosas batidas. Estas anémonas se fotografiaron en un canal de la isla Tatoosh, en aguas del extremo noroeste del estado de Washington. En un canal de marejada, estas grandes anémonas verdes, que son sésiles, se enfrentan al fuerte flujo situándose muy apretadas contra el fondo, donde el flujo es inferior al de la corriente principal. En un lugar más protegido, donde buena parte de la fuerza la absorben las rocas encaradas al mar, *A. xanthogrammica* alcanza mayor altura.

tran las anémonas. El flujo de la corriente principal puede alcanzar hasta los cinco metros por segundo en los días no tormentosos; a nivel de los tentáculos de las anémonas es mucho menor. Aunque hay ejemplares de anémonas verdes que llegan a los 20 centímetros de diámetro, en un canal expuesto adquieren forma de torta, con una altura media sobre la roca de 2,5 centímetros. En otras palabras, evitan el flujo rápido aplastándose contra el fondo.

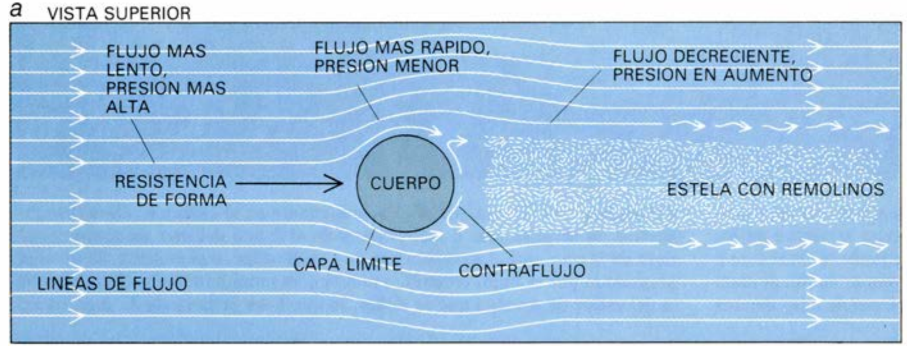
Algunas anémonas verdes pueden encontrarse asimismo en lugares más protegidos, donde el impacto de las olas cae sobre las rocas que dan al mar. En tales hábitats las anémonas miden, por término medio, siete centímetros de altura. Aunque el flujo de la corriente principal en lugares protegidos es mucho más lento que en los canales de marejada, las anémonas más altas se yerguen en el flujo y, por tanto, se encuentran con velocidades iguales o superiores a aquéllas con que tienen que habérselas sus parientes bajas de los canales expuestos.

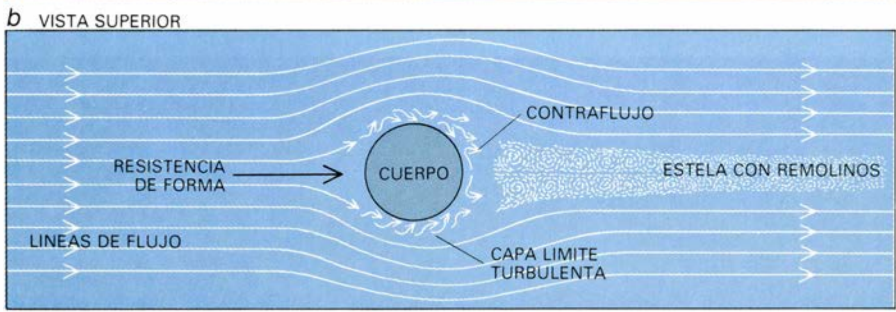
Las anémonas verdes de los lugares menos expuestos conviven, frecuentemente, con otras anémonas más pequeñas de la especie *Anthopleura elegantissima*, las llamadas anémonas colectivas, que se reproducen por división longitudinal en dos; los clones formados llegan a alfrombrar las rocas. Observé que el agua que bañaba a estas anémonas, pequeñas y estrechamente apretadas, viajaba a una velocidad diez veces menor que la del agua que bañaba una gran anémona verde solitaria situada entre las anteriores. El contraste indica cuán diferentes pueden ser los microhábitats de flujo de organismos que viven uno al lado de otro.

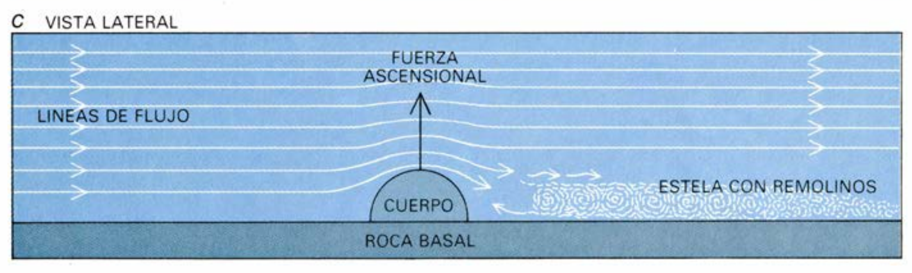
Muchos animales coloniales marinos, como los corales y los hidroideos, segregan un esqueleto que sostiene a los individuos de la colonia (los pólipos). Los animales crean el terreno donde se desarrollan y modifican así sus propios ambientes de flujo. Con el fin de visualizar el flujo alrededor de ramas de corales y de modelos artificiales de éstos, John Chamberlain, del Brooklyn College de la Universidad de Nueva York, y Richard Graus, de la Universidad de Cleveland, añadieron colorantes al agua. Como cabía esperar, observaron que el diámetro, el espaciamiento y la disposición de las ramas afectaban a la velocidad y dirección del movimiento del agua en la colonia. Pero encontraron también que configuraciones coloniales muy diferentes podían originar la misma pauta de flujo sobre los pólipos. Los pólipos de los látigos de mar y de los abanicos de mar (gorgonias) ocupan esqueletos flexibles que cimbrean, adelante y atrás, al paso de las olas. Puesto que el esqueleto se mueve con el agua, el flujo con respecto al pólipo puede ser bajo, aunque la colonia se agite a uno y otro lado en un hábitat de fuerte marejada.

La forma, el tamaño y la textura de un organismo sésil determinan la magnitud de las fuerzas mecánicas que se ejercen sobre él cuando el agua fluye a su través. Las fuerzas de arrastre o de resistencia al avance, que tienden a empujar al organismo corriente abajo, se deben a la viscosidad del agua, es decir, a la tendencia de las moléculas de agua a resistirse a resbalar unas sobre otras. El agua se pega a las superficies sólidas, incluidas las superficies de los organismos, y la capa límite de agua que rodea al organismo se halla sujeta a fuerzas de cizalla cuando el flujo pasa sobre éste. Puesto que el agua es viscosa y se resiste a dividirse, sobre el organismo se aplica una resistencia de fricción superficial. La magnitud de la resistencia es proporcional a la velocidad del flujo y a la longitud del objeto; y aumentará con el tamaño los organismos o al enfrentarse éstos a flujos más rápidos.

La resistencia que se ejerce sobre individuos pequeños en un flujo lento se debe principalmente a fricción superficial. Los grandes y sometidos a un flujo rápido sufren una resistencia adicional, la llamada resistencia de forma, que suele ser mucho más intensa que la fric-





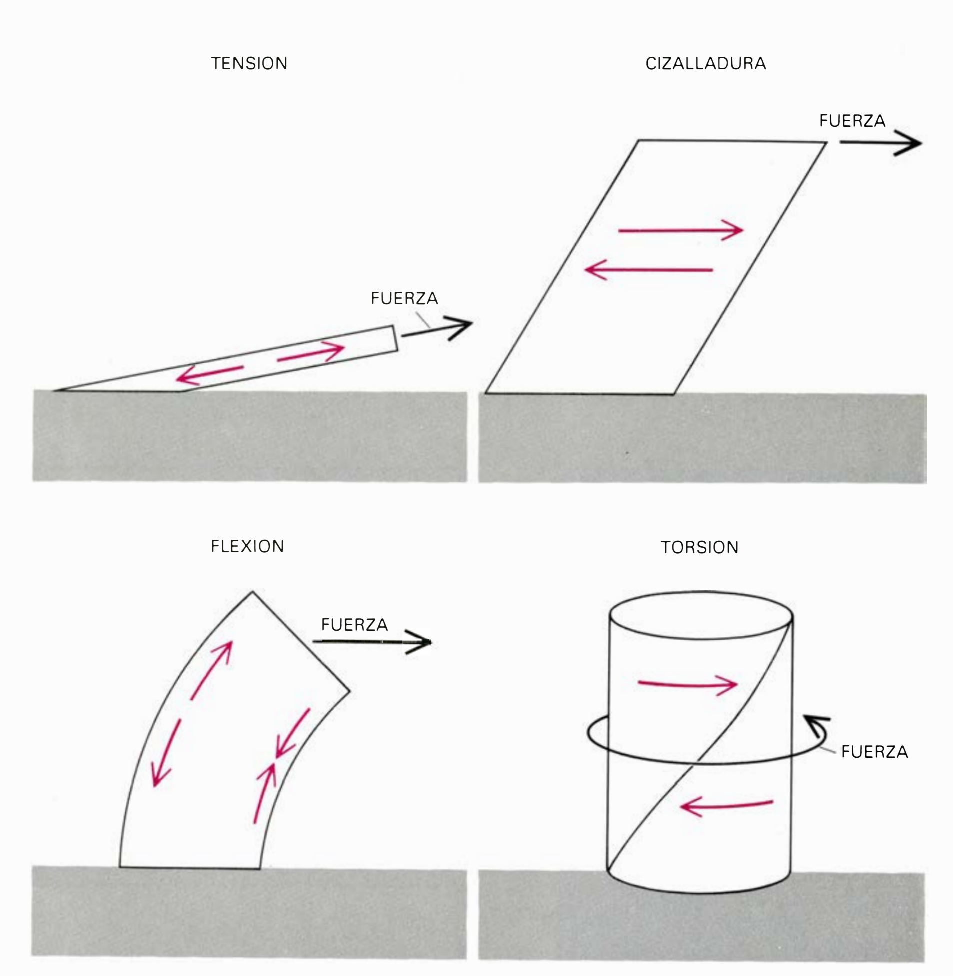


3. FUERZAS DE FLUJO que actúan sobre un organismo marino sésil, representadas esquemáticamente. Cuando la velocidad aumenta a lo largo de una línea de flujo (la trayectoria que sigue una partícula de fluido) la presión disminuye, y viceversa. Cuando el agua fluye alrededor de la sección más ancha de un cuerpo (*a*), su velocidad aumenta y la presión que ejerce sobre la superficie del cuerpo disminuye. Pasada la sección más amplia, el fluido se hace más lento y la presión aumenta. La viscosidad hace que las partículas de fluido de la capa límite (la que está más cerca del cuerpo) pierdan impulso cuando sobrepasan el cuerpo. Cuando el flujo sobre el cuerpo es suficientemente rápido y el cuerpo suficientemente grande, este retardo debido a la viscosidad frena el avance del agua en la capa límite, que entonces puede ser empujada corriente arriba por el aumento de presión del fluido que la rodea y que pierde aceleración. Detrás del cuerpo, una estela de fluido que se arremolina se separa del flujo principal, de modo que aquí la presión es menor que en la parte anterior. Esta presión neta corriente abajo se denomina resistencia de forma. Si el fluido se mueve con suficiente rapidez con respecto al cuerpo (*b*), la capa límite se hace turbulenta y existe una transferencia de impulso hacia ella procedente del flujo principal, lo que origina una estela menor y menos resistencia de forma. Cuando el fluido se desplaza más deprisa por la parte superior o a un lado del cuerpo (*c*), la disminución de presión succiona el cuerpo. Se dice que esta fuerza es ascensional, pero no siempre es hacia arriba; puede ejercerse en cualquier dirección ortogonal al flujo.

ción superficial. En el lado del organismo que queda corriente abajo se forma una estela turbulenta. La presión que se aplica sobre el lado situado corriente arriba es mayor que la del lado situado corriente abajo, por lo que el organismo recibirá un empuje en ese sentido. Cualquier característica de un organismo que reduzca el tamaño de la estela relajará la resistencia de forma. La magnitud de la resistencia de forma es proporcional a la superficie del cuerpo y al cuadrado de la velocidad: un aumento relativamente pequeño en la longitud o en la velocidad puede llevar a un aumento comparativamente grande en la resistencia de forma.

El investigador que pretende estudiar de qué modo la forma, la textura superficial y la flexibilidad de un organismo afectan a las fuerzas de resistencia puede hacerlo siguiendo distintos caminos. Pueden compararse las fuerzas de resistencia y las pautas de flujo alrededor de organismos similares que difieran en características particulares. Pueden hacerse asimismo modelos de los organismos, modificando varias características con el fin de medir su efecto sobre la resistencia. Se fijan organismos sésiles o modelos de los mismos a transductores de fuerza en la costa o en un tanque de flujo o una instalación de remolque, en los que se controla la velocidad de flujo. Las pautas del flujo alrededor de los cuerpos se hacen visibles mediante colorantes o partículas marcadoras, o bien se miden por medio de pequeñas sondas electrónicas.

Si un organismo grande tiene la mayor parte de su superficie en paralelo al flujo, en lugar de tenerla en ángulo recto con respecto a éste, creará una estela menor y con ello reducirá la resistencia de forma. Esta manera simple de reducir la resistencia está bien ilustrada por las fuerzas que actúan sobre dos especies de anémonas de mar grandes: la verde aplanada, de que me ocupé antes, y *Metridium senile*, una especie alta y plumosa que tiene una corona amplia y ondulada de diminutos tentáculos que actúa de filtro para capturar zooplancton. La anémona plumosa vive en la zona submareal, donde las lentas corrientes de marea inclinan su corona en ángulos rectos con respecto al flujo. Por el contrario, la anémona verde sitúa la mayor parte de su superficie paralela al flujo. A una velocidad determinada, tanto el tamaño de la estela como la cantidad de resistencia son mayores en una anémona plumosa que en una verde del mismo diámetro coronal.

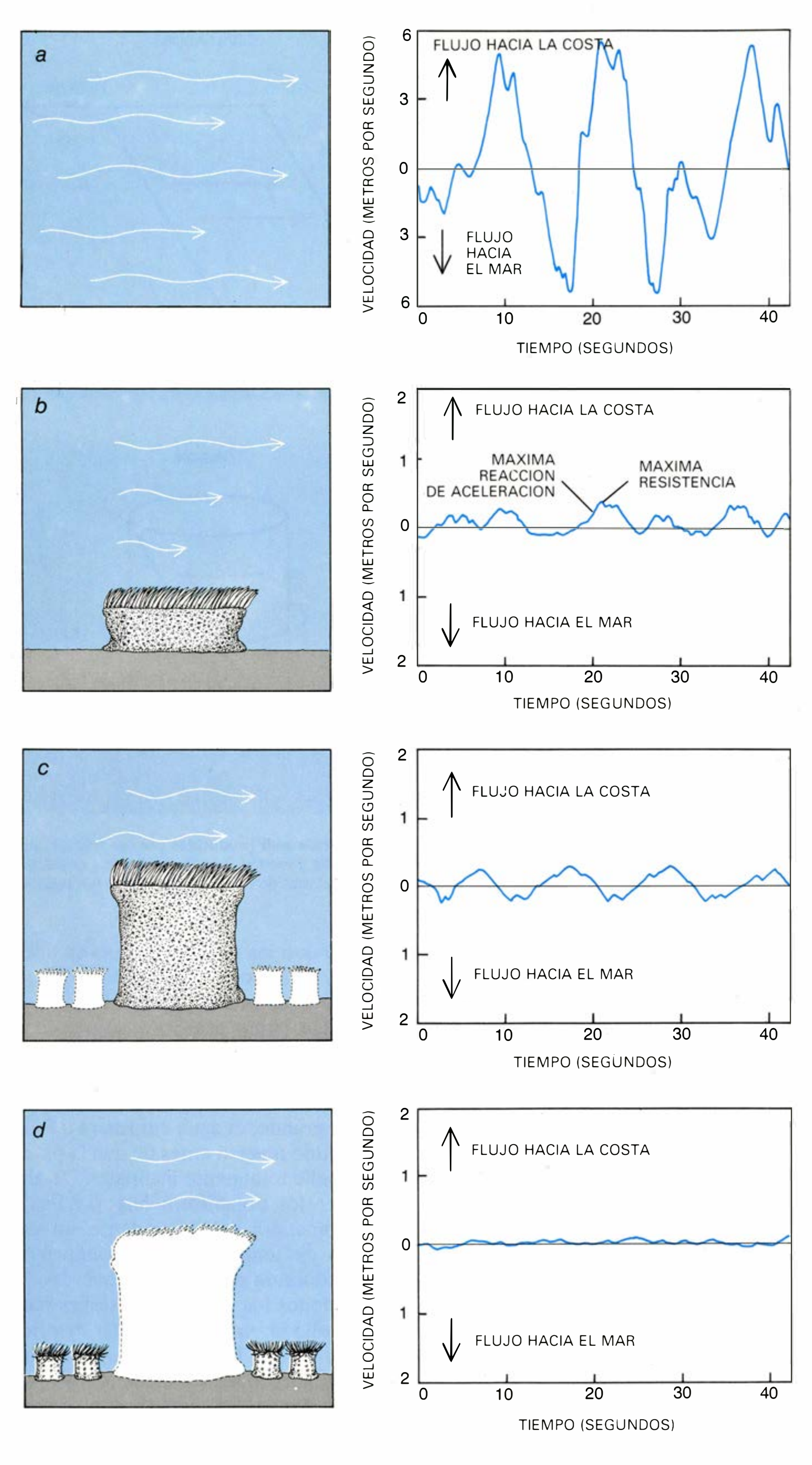


4. EFECTOS SOBRE LOS TEJIDOS de un organismo marino sésil producidos por las fuerzas que genera una corriente de agua; son: tensión, cizalladura, flexión y torsión. Las flechas negras muestran aquí la carga que se ejerce sobre el organismo; las de color, el sentido en que se deforman sus tejidos.

Las ramas y los tentáculos de muchas plantas y animales marinos flexibles se agrupan cuando el organismo se inclina por la acción del agua que fluye. El movimiento sitúa al organismo en una posición de mayor paralelismo con el flujo, relajando la resistencia y achicando el tamaño de la estela. Vale de ejemplo la anémona plumosa: cuando el agua fluye más deprisa, la gran corona tentacular se invierte como un paraguas abierto al que el viento da la vuelta. La comparación entre la resistencia que ejerce un modelo de anémona plumosa con corona rígida y la resistencia de un modelo con corona flexible revela que el cambio de forma pasivo de la anémona viva reduce, de manera significativa, la resistencia.

Las fuerzas de resistencia que actúan sobre muchas algas grandes de costas batidas por el oleaje son sorprendentemente reducidas, a pesar del enorme tamaño de algunos vegetales de éstos. Debido a su flexibilidad, las plantas pueden inclinarse paralelamente al flujo y situarse más cerca del fondo, donde el flujo es más lento. La flexibilidad les permite a las algas emplear otra maniobra de elusión ante el oleaje. Puesto que las plantas se mueven adelante y atrás con el agua que oscila, la corriente no es muy grande hasta que se hallan completamente desplegadas en el sentido del flujo. Cuanto más larga es el alga, más lejos la lleva el agua en su avance. Si un alga es lo suficientemente grande, el agua empezará a fluir en sentido inverso antes de que la planta se halle totalmente inclinada. De ahí que, en los organismos largos y flexibles sometidos a flujo oscilante, un aumento de longitud puede conducir a una reducción de la resistencia.

No todos los organismos sésiles han desarrollado una morfología externa que minimiza la resistencia. Algunos corales gorgonarios y alcionáceos producen colonias en abanico que se yerguen en ángulo recto con respecto a la dirección del flujo. Estos animales coloniales se alimentan de materiales suspendidos en el agua. La orientación de la colonia en un plano perpendicular al flujo significa no sólo que ésta se halla expuesta a la máxima cantidad de flujo de agua por unidad de tiempo, sino también que no quedan miembros corriente abajo, que, de otro modo, tendrían que volver a procesar agua de la



5. VARIACIONES DEL FLUJO alrededor de anémonas sésiles. El flujo de la corriente principal en un canal de marejada (*a*) es con frecuencia de hasta cinco metros por segundo al romper las olas en la costa y, luego, retroceder el agua. Una anémona verde en un canal de este tipo (*b*) adopta una forma aplanada para aprovechar la menor velocidad del flujo cercano al fondo del canal. La máxima fuerza de resistencia sobre la anémona se ejerce cuando la velocidad del agua es mayor, mientras que la fuerza aceleración-reacción máxima se ejerce cuando la aceleración es mayor. (La aceleración es elevada cuando la pendiente del trazado de la velocidad con respecto al tiempo es empinada.) En un lugar protegido (*c*), donde el impacto del oleaje rebota en las rocas hacia el mar, una anémona verde penetra más en el flujo. El flujo que la rodea es muy parecido al que rodea a la anémona del canal de marejada. Algunas de las anémonas colectivas (*d*) situadas cerca de la anémona verde del lugar protegido sienten todavía menos flujo.

que ya se han extraído las partículas alimenticias. Cuando trabajaba en el laboratorio marino de la Universidad de Duke, Gordon Leversee demostró que los corales gorgonarios planos capturaban más alimento si se encontraban perpendiculares al flujo que cuando se hallaban dispuestos en paralelo.

¿De qué modo consiguen la orientación correcta las gorgonias en abanico? Stephen A. Wainwright y John Dillon, de la Universidad de Duke, observaron que las gorgonias jóvenes se orientaban en todos los sentidos. Sugieren ambos que, cuando el agua en movimiento ejerce presión sobre los abanicos en crecimiento, tiende a girarlos hasta que se hallan orientados perpendicularmente al flujo.

Los organismos marinos sésiles inmersos en agua en movimiento pueden hallarse asimismo sometidos a fuerzas ascensionales. Al curvarse hacia arriba la superficie apical del organismo, el agua que fluye sobre la misma debe desplazarse más rápidamente que el agua que fluye sobre otras superficies. La presión en la región de flujo rápido es inferior a la que se ejerce sobre otras regiones y, así, se genera una fuerza que tiende a levantar del sustrato al organismo. Mark W. Denny, de la Universidad de Stanford, ha medido fuerzas ascensionales significativas en varios organismos, como bellotas de mar y lapas, que se adhieren a las rocas. Por el mismo motivo, si la forma de un organismo es tal que el agua fluye más deprisa alrededor de uno de sus lados que en torno al otro, el individuo se verá arrastrado hacia el lado donde el flujo es más rápido.

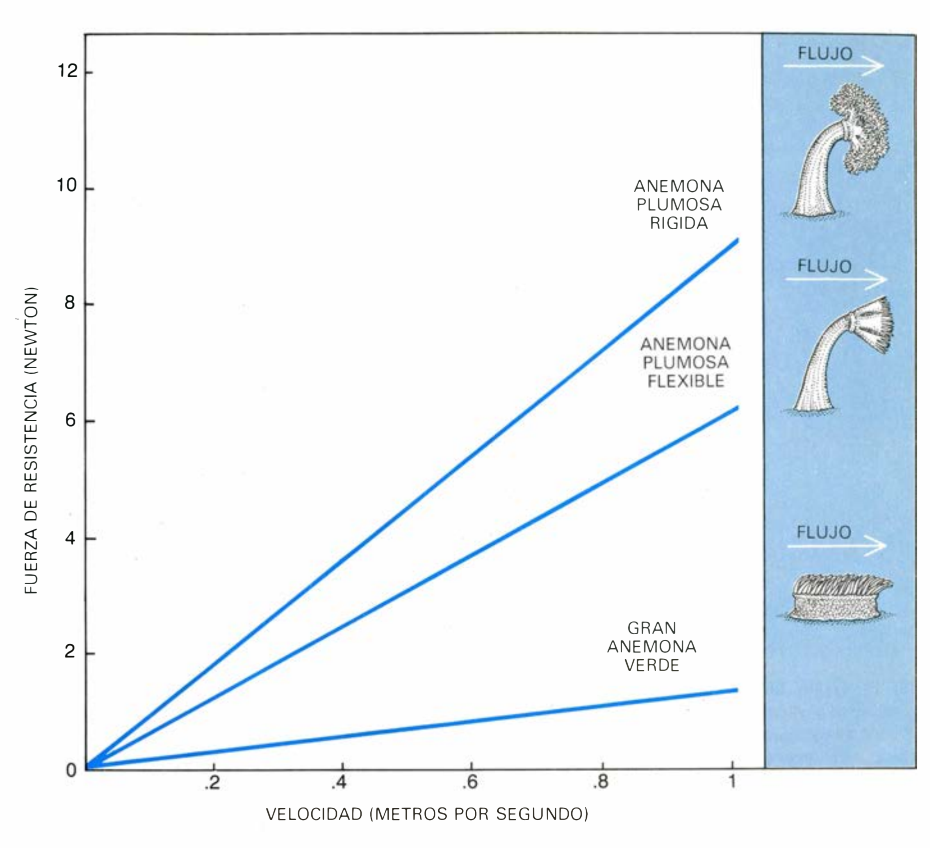
Otra fuerza que actúa también en muchos organismos sésiles es la reacción de aceleración. El agua que rodea a una planta o a un animal sobre el que pasan olas de forma regular aumenta su velocidad, la reduce e invierte el sentido. La reacción de aceleración empuja al organismo en el sentido en que el agua se acelera. (Recordemos que cuando el agua reduce su velocidad está acelerando en sentido opuesto al de su movimiento.) Cuanto mayor es el volumen de agua que tiene que acelerarse para rodear un organismo, más fuerte es la reacción de aceleración. Un gráfico del flujo que incide sobre un organismo en una ola muestra que la resistencia máxima y la máxima reacción de aceleración no coinciden en el ciclo de onda. La fuerza que actúa sobre un organismo en un instante dado es la suma de la resistencia y la aceleración en ese instante.

Al examinar las ecuaciones que expresan la magnitud de las fuerzas de resistencia, de ascenso y de aceleración sobre los organismos sésiles, se aprecia que tanto la resistencia como la fuerza ascensional son proporcionales al volumen del organismo. Parecería, pues, que al aumentar un organismo de tamaño la reacción de aceleración resulta más importante que la resistencia o el ascenso. Denny y Thomas Daniel, de Duke, y el autor, proponen que las fuerzas de reacción-aceleración pueden establecer un límite máximo al tamaño que pueden alcanzar muchos organismos agitados por el oleaje. Ciertamente, muchos animales y plantas submareales, que viven en corrientes más lentas y uniformes que los organismos intermareales, alcanzan tamaños mucho mayores que éstos.

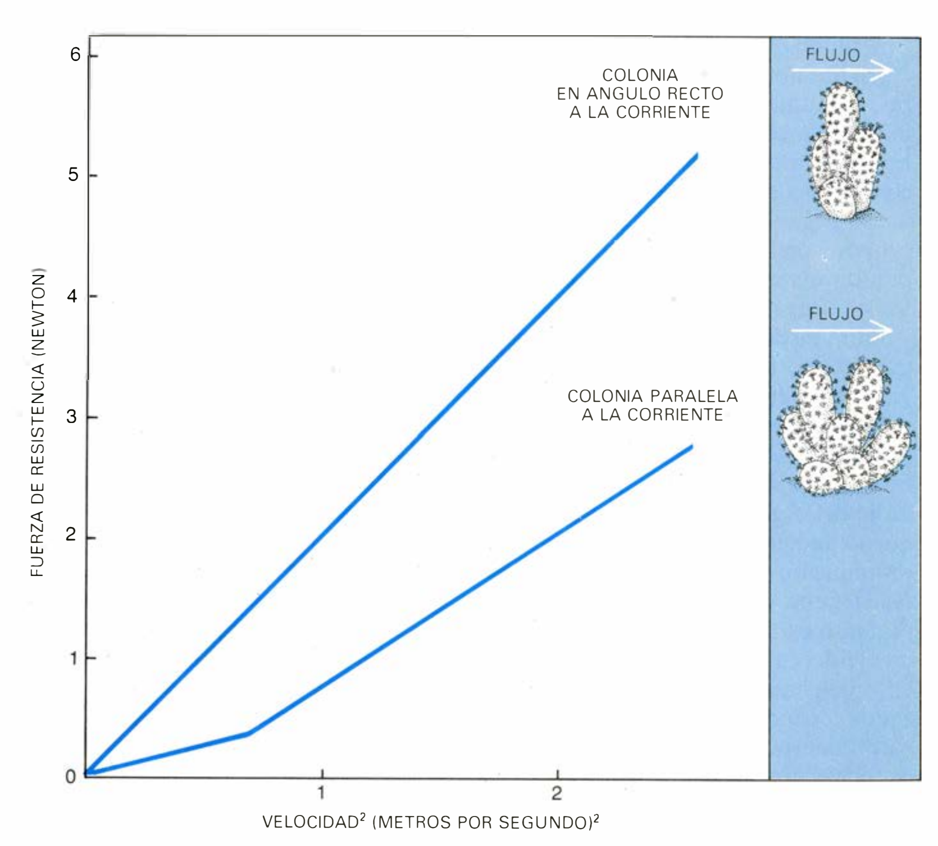
Quizá la acción del oleaje establezca indirectamente un límite superior al tamaño de determinados organismos al restringir el tiempo de que disponen los animales para comer. Suzanne Miller, de la Universidad de Washington, encontró que la fuerza necesaria para despegar a varios caracoles marinos de una roca era menor cuando éstos se movían que cuando se quedaban quietos. Si las olas que rompen acortan el tiempo que tiene un caracol para desplazarse y alimentarse, el crecimiento del animal puede verse limitado. De manera similar, el crecimiento de un organismo que se alimenta de partículas en suspensión en el agua puede verse limitado si un flujo violento le dificulta la captura de las partículas alimenticias o le obliga a retraer sus delicadas estructuras tróficas.

Las fuerzas inducidas por el flujo deforman o rompen a veces un organismo. Que ello ocurra o no depende de la magnitud de la tensión (fuerza por unidad de superficie) ejercida sobre los tejidos de un organismo situado en un flujo y de la respuesta de los tejidos a la tensión. El tamaño y la forma de un organismo determinan la magnitud de las tensiones que sufren sus tejidos cuando soportan una determinada carga, una fuerza de resistencia por ejemplo. Una planta o un animal fijos pueden soportar esa carga de varias maneras, entre ellas la tracción, el cizallamiento, la flexión y la torsión.

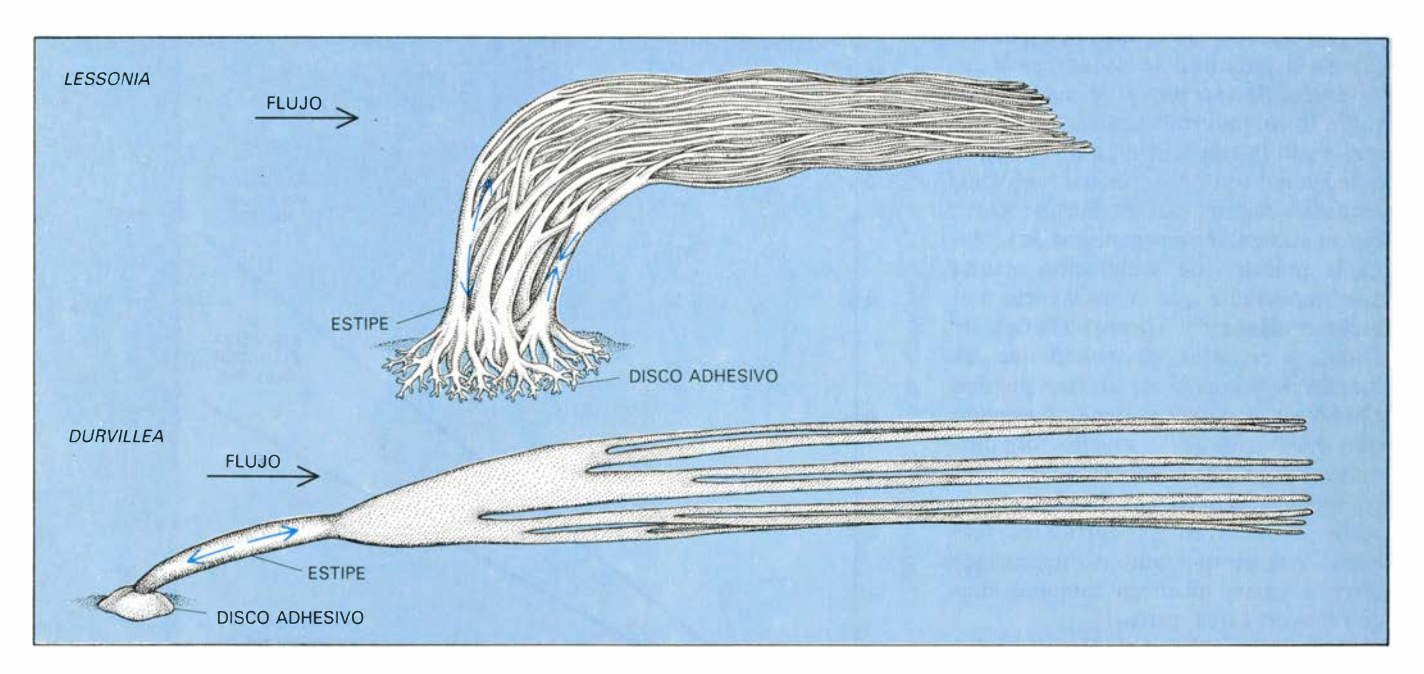
Cuando un organismo sufre cizallamiento o tracción, las tensiones que soportan sus tejidos son inversamente proporcionales al área de su sección transversal, pero no dependen de la longitud y la forma de esa sección transversal. Las tensiones ejercidas



6. INVESTIGACION DE MODELOS: las fuerzas de resistencia que actúan sobre las anémonas varían según la configuración que el animal presenta al flujo. La resistencia que se aplica sobre un modelo de anémona verde con la mayor parte de su superficie paralela al flujo es menor que la que ofrecen los modelos de la anémona plumosa. La anémona verde genera una estela menor, con lo que reduce la resistencia de forma. La resistencia de un modelo flexible de anémona plumosa que, como el animal vivo, adopta una forma más hidrodinámica al crecer la velocidad del flujo, es menor que la de uno rígido.



7. LAS FUERZAS DE RESISTENCIA que actúan sobre una colonia de coral varían según la orientación de la colonia en la corriente. El coral alcionáceo aquí ilustrado es *Alcyonium digitatum*, conocido también por mano de muerto, que forma una colonia carnosa que se parece algo a una mano abotagada.



8. TENSION SOBRE LOS TEJIDOS de un organismo marino sésil; puede reducirse si el organismo soporta la carga en tensión en lugar de flexionarse. Valen de ejemplo dos grandes algas que forman una densa banda a lo largo de las costas rocosas de Chile. *Lessonia nigrescens* se dobla ante el movimiento del agua, de manera que un lado de sus estipes se flexiona y el otro se comprime; los esfuerzos máximos, tensor y compresor, en el estipe son proporcionales a la longitud de éste e inversamente proporcionales al cubo de su radio. *Durvillea antarctica*, por el contrario, se deja arrastrar por el flujo de agua. El esfuerzo tensor que sufre su estipe depende de la longitud y es proporcional al cuadrado del radio. Las flechas de color representan los esfuerzos.

sobre los tejidos de un organismo estrecho son mayores que las que actúan sobre un organismo más ancho que soporte la misma carga.

Cuando un organismo se inclina por acción de un peso, un lado de su cuerpo se halla extendido y el otro encogido. Los tejidos de las superficies situadas corriente arriba y corriente abajo son, respectivamente, los que soportan una mayor tracción y los más comprimidos. La magnitud de la tensión tensional o compresiva en un punto del cuerpo de un tal organismo sésil es inversamente proporcional a la distancia entre ese punto y el extremo libre del organismo. Si dos organismos del mismo diámetro y altura distinta reciben la misma carga de flexión, las tensiones son mayores en el más alto.

La magnitud de las fuerzas tensional y de compresión asociadas a la flexión está en una relación inversa al cubo del radio del organismo. De ahí que un pequeño aumento del radio produzca una disminución importante de la tensión. Nada tiene de particular que muchos organismos que permanecen erectos en un fluido en movimiento sean más anchos y se hallen más fuertemente reforzados cerca de su base. Tampoco ha de sorprendernos que las constricciones estrechas que presentan los pedúnculos de muchos organismos sésiles funcionen a modo de articulaciones flexibles.

Un ligero aumento en el radio de un organismo que soporta un determinado par de torsión conduce también a un gran aumento de las tensiones torsionales de cizalladura. Por tanto, un organismo delgado se torcerá con mayor agilidad que otro grueso. Además, una cintura estrecha en el pedúnculo puede actuar como articulación rotacional.

Una muesca en el pedúnculo de una planta o de un animal fijado en un lugar azotado por la corriente puede resultar fatal para su vida; recuérdese el gran aumento de tensión que acompaña a la menor disminución del radio. *Nereocystis leutkeana*, una alga laminaria gigante del noroeste del Pacífico, proporciona un sorprendente ejemplo de cuán dañino sea el efecto de una tal depredación parcial. Entre el 30 y el 90 por ciento (según el lugar) de las plantas que llegan a la costa arrastradas por las olas han experimentado roturas en algún punto del estipe (pedúnculo), del que sólo falta una pequeña cantidad de tejido, por lo general roída por erizos de mar.

Las anémonas plumosas y las grandes anémonas verdes ilustran la importante relación que guardan la forma de un organismo y las tensiones a que se ve sometido. La fuerza del flujo sobre las anémonas plumosas de aguas calmadas y sobre las anémonas verdes de los canales de marejada es, en ambos casos, aproximadamente de un newton (unos 100 gramos) y tiende a empujar a los animales en el sentido de la corriente. Bajo una fuerza de esa magnitud, la tensión máxima que he calculado en los tejidos de una anémona plumosa, alta y delgada, es 45 veces superior, por término medio, a la que soporta una anémona verde, baja y ancha. En otras palabras, la anémona plumosa del hábitat "protegido" se halla en un ambiente que, desde el punto de vista mecánico, presenta muchas más tensiones que el hábitat "expuesto" de la verde.

Puede verse, por tanto, que una forma alta y delgada aumenta la tensión y una forma baja y ancha la reduce. Los organismos relajan la tensión de sus tejidos soportando cargas en tensión, en vez de flexionarse. Un ejemplo nos lo ofrecen dos grandes algas que forman un denso cinturón a lo largo de las costas rocosas de Chile: *Lessonia nigrescens* se mantiene erguida y se flexiona, en tanto que *Durvillea antarctica* se deja caer, doblándose por una constricción próxima al disco de fijación y es arrastrada por el agua en movimiento. Medí velocidades de flujo de hasta seis metros por segundo alrededor de ambas especies y fuerzas que rondaban los 20 newton. Un cálculo de la máxima tensión en un estipe de 20 centímetros de largo por dos de diámetro, sujeto a una carga de 10 newton, indica que ésta es de un tercio de newton por metro cuadrado si el estipe es arrastrado (como ocurre en *Durvillea*) y de más de 250 newton si el estipe se flexiona (así en *Lessonia*).

El grado de deformación de un organismo inmerso en agua que fluye depende, además de la magnitud de las tensiones que actúan sobre sus tejidos, de la rigidez de éstos. Se mide la rigidez tirando de un fragmento de tejido (en un aparato que se parece a un potro de tormento medieval) hasta que se rompe. El aparato mide la fuerza del tirón y la longitud que se ha distendido el tejido. La gráfica de los resultados revela que, para tirar de un tejido rígido hasta una determinada extensión, se necesita mayor tensión que para tirar de un tejido menos rígido. La pendiente de una curva tensión-extensión representa el módulo elástico de un material y nos mide la rigidez de éste.

Peculiaridad de muchos materiales biológicos flexibles es que su rigidez depende de la velocidad a la que se deforman (condición que no se da en el acero y el cristal, entre otros). Esta característica explica un rasgo interesante del diseño mecánico de los organismos sésiles. Como cabría esperar, un material se deforma antes bajo la acción de una tensión grande que bajo una tensión menor. La forma de un organismo determina la magnitud de la tensión en sus tejidos, y la magnitud de la tensión determina la tasa de deformación, que a su vez determina la rigidez propia de un material. Por consiguiente, la forma de un organismo puede afectar a la rigidez de sus tejidos.

La anémona de mar plumosa, *Metridium*, ejemplifica la incidencia del ritmo de vida en la rigidez de un organismo. Sometamos un tejido a una prueba de resbalamiento en la que se aplique una tensión constante y se mida la deformación durante un período de tiempo. Las anémonas de mar están compuestas básicamente de un tejido gelatinoso denominado mesoglea. Si exponemos en una gráfica los resultados de una prueba de resbalamiento para la mesoglea de *Metridium*, observaremos que el tejido se deforma poco cuando se le estira a lo largo de un intervalo inferior al minuto. Nos movemos todavía en la escala temporal de los cambios de forma musculares que desarrolla el individuo. De ahí que la mesoglea proporcione un apoyo razonablemente rígido para oponerse a los músculos y los elementos esqueléticos del organismo.

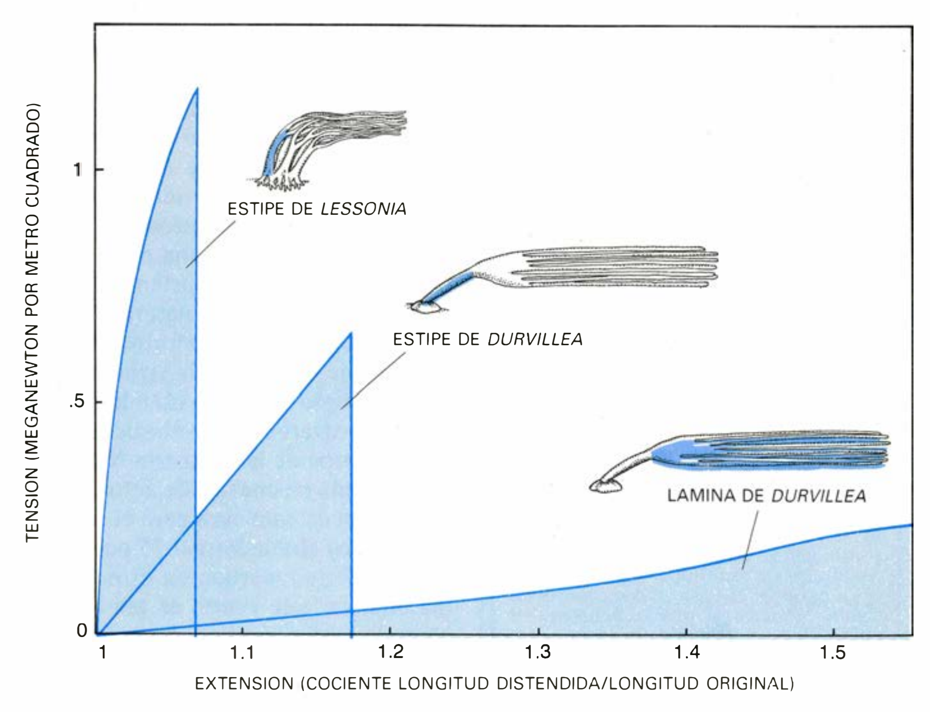
Considérese ahora una tensión aplicada durante varias horas, lapso de tiempo propio de las corrientes de marea que actúan sobre la anémona. La mesoglea se distiende de forma creciente y el individuo se inclina pasivamente en la dirección del flujo. A lo largo de un período de 12 a 24 horas, la mesoglea se alargará hasta tres veces su longitud original, incluso mediante tensiones pequeñas. Ese dilatado período corresponde al tiempo que una anémona plumosa invierte en hincharse hasta los grandes tamaños que suele alcanzar. (Lo hace absorbiendo agua a través de surcos denominados sifonoglifos, a cuyo través el agua se bombea mediante cilios batientes.) Para un músculo que trabaje rápidamente, la mesoglea es bastante rígida; en cambio, para un sifonoglifo que trabaja lentamente resulta bastante dúctil.

La curva de resbalamiento para un fragmento de mesoglea de una anémona verde es notablemente distinta. Aunque reciba durante todo el día los embates del oleaje en un canal de marejada, esta anémona no se distiende gran cosa. La mesoglea de las dos especies muestra discrepancias estructurales y moleculares que reflejan adaptaciones a las diferentes condiciones mecánicas de sus hábitats.

El pedúnculo de sostén de muchas plantas y animales fijados consta de más de un tipo de material. Si el material más rígido se halla cerca del centro del pedúnculo, la estructura es más flexible de lo que sería si el material se hallara en la periferia. Stephen Wainwright y el autor estudiaron un ejemplo de una estructura tan flexible que fácilmente se la puede anudar sin deformarla. Se trata del estipe de la laminaria *Nereocystis*. El tejido situado cerca del centro del estipe es más rígido que el tejido de la periferia. Ello no sólo sitúa al principal tejido que soporta tensiones donde es menos probable que sea ronzado por los erizos de mar, sino que, a la vez, proporciona a la laminaria la flexibilidad que le permite colocarse paralela a una corriente rápida, reduciendo así la resistencia. Por el contrario, los tallos de muchas plantas terrestres que se mantienen erguidas (los girasoles, por ejemplo) tienen vasos rígidos en su periferia.

La cuantía de la deformación a que un organismo sésil se somete ante el flujo de agua, en un determinado hábitat, importa sobremanera a la hora de asegurarse el éxito de sus actividades diarias. ¿Puede mantener sus tentáculos en la disposición adecuada para capturar partículas alimenticias? ¿Puede sostener sus superficies fotosintetizadoras por encima de las de sus competidores? ¿Corre el riesgo de ser dañado o arrancado de su sustrato por la corriente?

Que un organismo sésil sea arrancado de su sustrato y arrastrado por la corriente no sólo depende de la magnitud de las tensiones con las que se enfrenta;



9. RESISTENCIA DEL TEJIDO, que puede medirse calculando el área que queda bajo una curva de tensión-extensión de un determinado tejido que se extiende hasta romperlo. La superficie representa el trabajo por volumen de material que se requiere para romper el ejemplar. Los cálculos revelan que no hay diferencia significativa en la cantidad de trabajo (74 kilojoule por metro cúbico) que se precisa para romper los tejidos del fuerte estipe de *Lessonia*, el estipe más elástico de *Durvillea* o la lámina extensible de *Durvillea*. En los organismos marinos sésiles, pues, hay más de una manera de ofrecer resistencia.

tiene que ver también con la dureza y resistencia de sus tejidos y del pegamento que lo sujeta al fondo. La dureza de un material se define como la tensión necesaria para romperlo. Sin embargo, la probabilidad de rotura por una ola o una corriente dependerá de la resistencia que oponga, es decir, del trabajo que se requiere para romperlo. El cristal puede ser más duro que el cuero, pero es más difícil romper un zapato que un jarrón porque el cuero es más resistente que el cristal. El área que abarca una curva tensión-extensión de un individuo que se ha distendido hasta su fractura representa el trabajo por volumen de material necesario para conseguir la rotura; de otra manera: mide la resistencia del material.

Una serie de tales curvas revela que la fortaleza o resistencia puede lograrse siguiendo varias estrategias. Por ejemplo, el material rígido y resistente del estipe del alga *Lessonia* no es más fuerte que el estipe más dúctil del alga *Durvillea*. Estas dos algas ilustran dos estrategias muy distintas de resistencia a la rotura. Una es ser rígido y fuerte, como *Lessonia*; la otra es ser un "enclenque" y deformarse bajo una tensión, pero poseer al propio tiempo la capacidad de extenderse mucho antes de romperse, a la manera de *Durvillea*.

Si se aplica tensión a un enclenque durante un tiempo suficientemente largo hasta llegar al punto de rotura, se fracturará a una tensión menor que la que requiere otro organismo que sea rígido y duro. Por ello, el modelo de resistencia del enclenque es más eficaz para organismos sometidos a fuerzas pulsantes de corta duración, las que se dan, por ejemplo, en una costa batida por las olas. Pero si el enclenque quiere sobrevivir necesita una naturaleza elástica: poder recuperar su forma de reposo antes de que llegue la siguiente ola. Muchas algas que se extienden gozan de una extraordinaria elasticidad. El propio estipe de la laminaria *Nereocystis* almacena en energía de deformación (que después empleará para el retroceso elástico) alrededor del 75 por ciento de la energía invertida en estirarla. El material de este estipe es tan elástico como la misma elastina, la proteína semejante al caucho merced a la cual las arterias humanas restablecen rápidamente la forma tras cada pulso de sangre bombeada por el corazón.

He diseñado pruebas mecánicas para simular las tensiones a que se ven sometidos los organismos marinos sésiles en la naturaleza. Si se comparan con las bellotas de mar y con los corales pétreos, las anémonas de mar se agruparían en la clase de los enclenques; no obstante, mis pruebas de simulación con mesoglea de la anémona verde de un canal de marejada indican que este tejido recupera su forma de reposo después de cada ola. La mesoglea de la anémona plumosa de aguas calmas no se recupera tan deprisa bajo la misma tensión, y luego de muchos ciclos puede haberse distendido tanto que se rompe. La anémona plumosa es un organismo de la clase de los enclenques que aparentemente carece de la elasticidad necesaria para sobrevivir en un canal de marejada.

Varios ecólogos marinos se han ocupado del papel destacado que las perturbaciones físicas desempeñan en la diversidad de las comunidades de organismos sésiles. Los retazos de sustrato que quedan vacíos cuando se arrancan los organismos allí instalados pueden ser colonizados por otras plantas y animales que, sin esa alternativa, se verían excluidos del hábitat por organismos con más éxito competitivo. Quienes estudiamos los arrecifes de coral de Jamaica antes y después de que fueran asolados por el huracán Allen, en 1980, tenemos una vívida imagen del papel que el movimiento violento del agua puede desempeñar en la alteración de una comunidad de organismos. Como imaginará el lector después de mis comentarios sobre las características físicas de los organismos marinos sésiles, los corales, altos y rígidos, de los arrecifes resultaron aplastados por el huracán, mientras que las gorgonias, delicadas y flexibles, y las algas cespitosas bajas sobrevivieron (salvo las que fueron arrancadas, o quedaron enterradas por los cascotes de coral).

Se pueden citar innumerables ejemplos de las fantásticas y maravillosas características que han desarrollado los organismos marinos para resistir y sacar partido de los movimientos del agua. Los que he citado deberían ilustrar el hecho de que el biólogo marino, provisto de unas cuantas reglas físicas básicas, puede hacer progresos considerables en la comprensión de las proezas físicas de organismos de distinta construcción en diferentes hábitats. Un campo atractivo cuyo estudio apenas si ha empezado es el del papel que las fuerzas mecánicas cumplen en el crecimiento y desarrollo de los organismos, quizás ayudando a modelar su cuerpo en las formas adultas, tan bellamente diseñadas, que decoran los litorales y los arrecifes coralinos.

A la búsqueda de números primos

Hasta anteayer las pruebas para determinar si un número de 100 cifras era primo habrían durado un siglo, incluso con la ayuda de un computador. Hoy basta con un minuto

Carl Pomerance

Los números primos son unidades con las que se forman, por multiplicación, todos los demás. Si un número es primo, entonces no existen números naturales menores cuyo producto sea dicho número. Por ejemplo, 11 es un número primo, ya que no se puede expresar como producto de factores menores; sólo 1 × 11 es igual a 11. Por contra, si un número es compuesto, entonces se podrá expresar como producto de dos o más factores primos. Así, 12 es igual a 2 × 2 × 3. Todo número entero mayor que 1 es primo o producto de un conjunto único de números primos. Este hecho, conocido ya por los griegos de la antigüedad, es esencial para el sistema de los números naturales, hasta el punto de que se ha convenido en llamarle teorema fundamental de la aritmética.

¿Cómo determinar si cierto número es primo o compuesto? El procedimiento más directo consiste en dividir el número en cuestión por 2, 3, 4, y así sucesivamente. Si una de las divisiones es exacta (esto es, el resto de la misma es cero), entonces se tratará de un número compuesto, siendo el divisor y el cociente factores suyos. Pero si ninguna de las divisiones es exacta, para todos los números hasta el número en cuestión, nos hallaremos entonces ante un número primo. La verdad es que no se necesita llegar hasta ese número; el proceso termina en cuanto el divisor supera la raíz cuadrada del número. La razón estriba en que los factores siempre se presentan a pares; si uno de ellos es superior a la raíz cuadrada, entonces debe existir otro inferior a la misma.

La finalización de las divisiones al llegar a la raíz cuadrada hace que la prueba de primalidad sea mucho más rápida. Hay otros atajos, como el de borrar todos los números pares posteriores al 2. Sin embargo, este método de división es totalmente inviable cuando se trata de los mayores números primos que se conocen. Considérese por ejemplo el número $2^{44.497} - 1$, que tiene 13.395 cifras. Este número es primo, como lo demostraron, en 1979, Harry L. Nelson y David Slowinski, del Laboratorio Lawrence Livermore. Si un computador llevase a cabo las divisiones a razón de un millón por segundo, y si parara al llegar a la raíz cuadrada del número, emplearía 10^{6684} años en efectuar el trabajo.

La dificultad con el método expuesto está en que hace mucho más de lo requerido: no sólo decide si un número es primo o compuesto, sino que también proporciona factores de cualquier número compuesto. Aunque existen métodos de factorización que no se basan en divisiones de prueba, ninguno de ellos puede, en un tiempo razonable, ni siquiera con un gran computador, factorizar un número arbitrario con "tan sólo" 100 cifras. Ahora bien, resulta que es posible determinar si un número es primo sin necesidad de hallar, en el caso de que sea compuesto, factores del mismo. Si el número no posee factores pequeños, estos métodos son invariablemente más eficientes que los métodos que determinan factores. A lo largo de los dos últimos años se ha desarrollado un método que permite determinar la primalidad de un número de 100 cifras, por medio de un computador, en unos 40 segundos.

El problema de los criterios de primalidad, y el superficialmente relacionado de la factorización de un número, son clásicos en la teoría de números, la rama de la matemática que se ocupa de las propiedades de los números enteros. La teoría de números está repleta de problemas tentadoramente sencillos de enunciar, pero notoriamente difíciles de resolver. Los problemas de la teoría de números relacionados con la primalidad han constituido un acicate y un reto para los matemáticos desde los tiempos de Euclides, por lo menos.

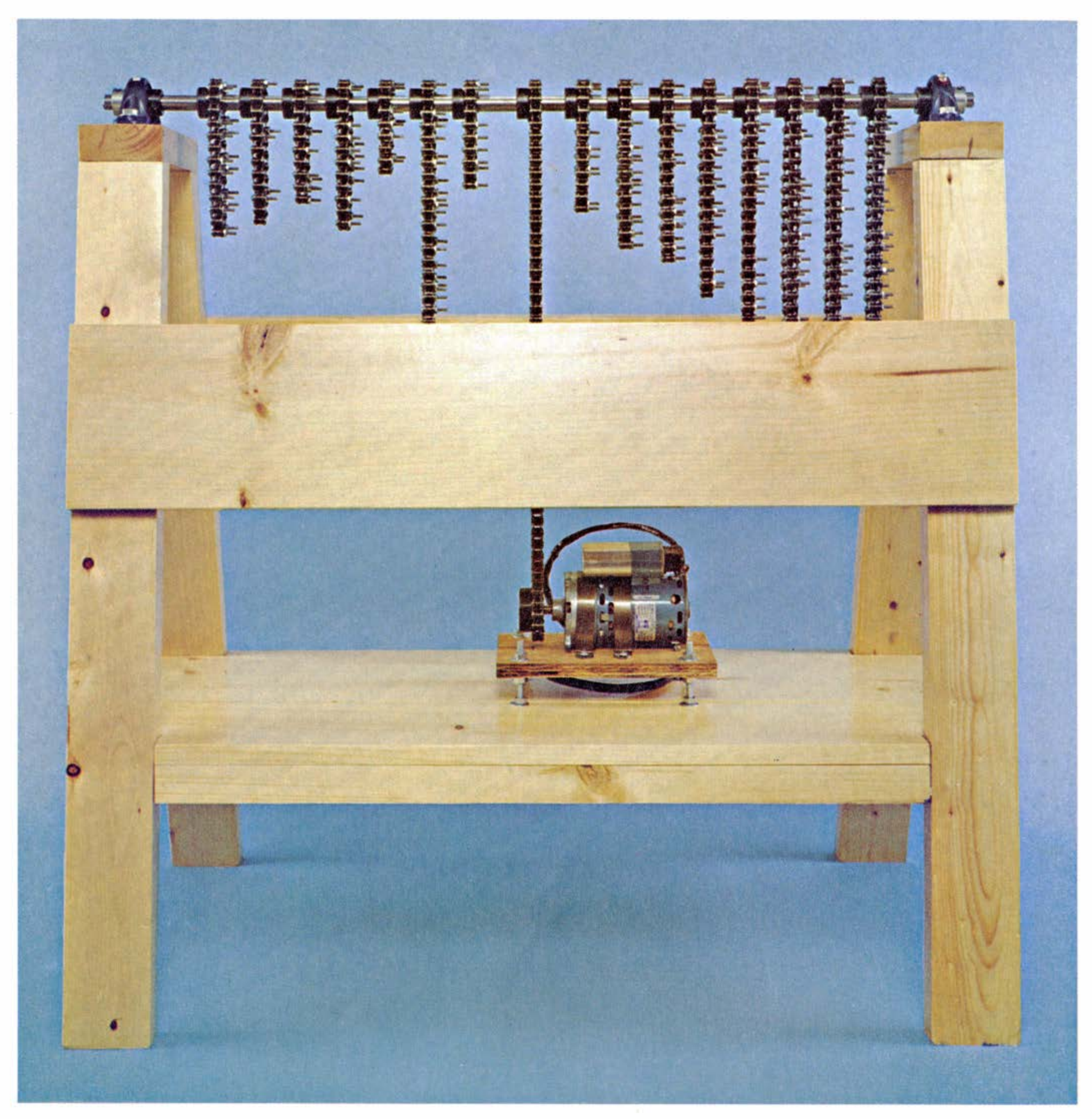
Hay, al parecer, infinitos pares de números primos mellizos, esto es, pares de números primos como 17 y 19 cuya diferencia es 2, conjetura que nunca se ha logrado demostrar. Tampoco se ha probado el enunciado, muy verosímil, de que existe al menos un número primo entre dos cuadrados perfectos consecutivos. En 1742 Christian Goldbach conjeturó que todo número par mayor que 2 era la suma de dos primos. Así, por ejemplo, 32 es la suma de 13 y 19. La conjetura de Goldbach también ha resistido todos los intentos de demostración, aunque en 1937 el matemático ruso I. M. Vinogradov mostró que todos los números impares "suficientemente grandes" se podían expresar como suma de tres números impares, aunque no logró explicar qué debía entenderse con la expresión "suficientemente grande".

Un corolario del teorema de Vinogradov es que todo número par suficientemente grande se puede expresar como la suma de 4 números primos, ya que, de acuerdo con el teorema, siempre es posible representar el número impar que resulta de sustraer 3 a un número grande como una suma de tres primos, por lo que dicho número es la suma de los tres números primos más el número primo 3. En 1966, el matemático chino Chen Jing-run mostró que todos los números pares suficientemente grandes se podían expresar como la suma de un número primo y un número que fuera primo o el producto de dos números primos. Estas aproximaciones a la conjetura de Goldbach son resultados profundos, en el sentido de que sus demostraciones, que presentan un alto grado de dificultad, utilizan análisis matemático avanzado.

Abundan también los enunciados sobre números primos que admiten una demostración elemental, y hay pruebas deliciosamente ingeniosas. Por ejemplo, desde Euclides se sabe que el conjunto de los números primos es infi-

nito. El argumento es indirecto. Si se supone que el conjunto de los números primos es finito, se llega a una contradicción. En efecto, sea p el mayor número primo, que existiría en el caso de que el conjunto de los números primos fuese finito. A continuación consideremos el número N, definido como el producto de todos los números primos desde 2 hasta p. El número $N + 1$ será o bien primo o bien compuesto. Como $N + 1$ es mayor que p, por la suposición inicial será compuesto, pues de lo contrario sería un primo superior a p. Por ser $N + 1$ compuesto, del teorema fundamental de la aritmética deducimos que posee factores primos. Pero a causa del modo como se ha definido $N + 1$ ocurre que el resto de la división de $N + 1$ por cualquier primo desde 2 hasta p es igual a 1. Así que sus factores primos serían mayores que p. Por consiguiente, no puede existir un límite superior al conjunto de los números primos, ya que esta suposición lleva a una contradicción.

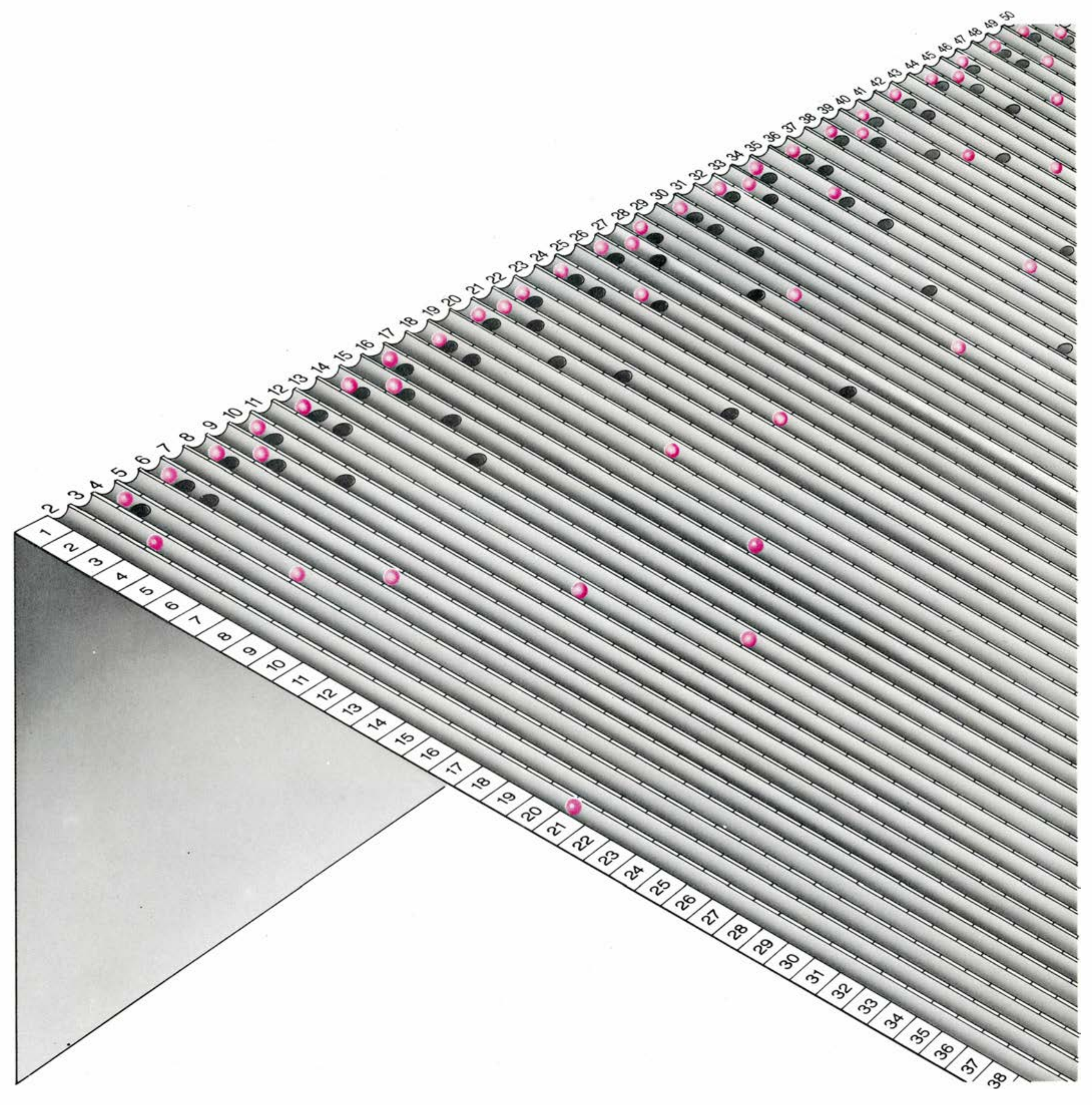
De una forma similar es fácil demos-



1. MAQUINA PRIMITIVA para la exploración del sistema de los números construida en 1926 por D. H. Lehmer, de la Universidad de California en Berkeley. Se montó con materiales de fácil obtención, como un caballete de aserrar y cadenas de bicicleta. Era un computador con una finalidad especial: se programaba al objeto de hallar prestamente números de una forma especial requerida para la solución de ciertos problemas en teoría de números. Uno de los problemas más importantes entre éstos era el de los criterios de primalidad: decidir si un número es primo o compuesto. (Un número primo sólo es divisible por sí mismo y por 1; si tiene otros divisores es compuesto.) Las condiciones que debe satisfacer una solución numérica de un problema se podían programar en la máquina de Lehmer introduciendo pasadores en ciertas anillas de las cadenas. Estas, impulsadas por un motor, giraban hasta que todos los pasadores quedaban alineados, momento en que el motor se paraba automáticamente. El número correspondiente a la configuración de las cadenas al parar satisfacía las condiciones programadas. Lehmer construyó versiones más rápidas, pero el ejemplar de 1926 se destruyó. La máquina de la fotografía la está reconstruyendo Roberto Canepa, de la Universidad de Carnegie-Mellon, en el Museo del Computador de Marlboro, Massachusetts.

trar la existencia de números primos consecutivos tan distantes unos de otros cuanto se quiera. Consideremos la sucesión de números $n! + 2$, $n! + 3$, $n! + 4$, ..., $n! + n$, donde $n!$ (léase n factorial) es el producto de todos los números enteros desde 1 hasta n. Obsérvese que $n! + 2$ es divisible por 2, que $n! + 3$ es divisible por 3 y así sucesivamente hasta llegar a $n! + n$, que es divisible por n. Todos los $n - 1$ números de la sucesión serán, pues, compuestos. Podemos imaginar una sucesión de la longitud que nos apetezca: basta tomar n suficientemente grande.

Muchos matemáticos han considerado la teoría de números como "la reina de la matemática", en parte por la intrincada belleza de sus demostraciones, pero también porque ha perdurado el sentimiento de que su estudio es una forma de contemplación pura sin el lastre de potenciales aplicaciones prácticas. Sin embargo, el desarrollo de la teoría de números ha venido estimulado también, desde 1977, por el reconocimiento de que quizás estuviera llamada a cumplir una importante misión en criptografía, el estudio de la seguridad en las comunicaciones. En dicho año



2. CRIBA PARA NUMEROS PRIMOS, atribuida el sabio griego Eratóstenes, uno de los primeros métodos inventados para distinguir los números primos de los compuestos hasta un cierto número predeterminado. La criba de Erastótenes está representada en la ilustración por un plano inclinado con agujeros. Los números en cuestión se hallan simbolizados por bolas que caen rodando por surcos del plano. Los agujeros se distribuyen de la siguiente manera: En primer lugar, los de la segunda fila están situados en columnas alternas empezando por 4, de modo que todas las bolas con número par, excepto 2, se deslizan por los agujeros. Se busca a continuación el surco sin agujeros encabezado por el menor número, en este caso el 3, y se practican agujeros en un surco de cada 3, excepto el 3, en la tercera fila. El procedimiento prosigue perforando agujeros en un surco de cada 5 en la quinta fila, de cada 7 en la séptima, y así sucesivamente hasta llegar a la fila cuyo número es el mayor que no supera la raíz cuadrada del número mayor encima de los surcos. Las bolas que no caen por algún agujero corresponden a los números primos. Por ejemplo, los 25 números primos inferiores a 100 se pueden determinar recogiendo las bolas, con números entre 2 y 100, que sobrepasan la séptima fila. (7 es el mayor número que no supera la raíz cuadrada de 100.)

Ronald L. Rivest, del Instituto de Tecnología de Massachusetts, Adi Shamir, del Instituto Weizmann de Ciencias, y Leonard M. Adleman, de la Universidad de Southern California, señalaron que se podría basar un sistema criptográfico de clave pública en la dificultad de factorizar un número compuesto que sea el producto de dos números primos de cien cifras cada uno.

En un sistema de clave pública, los medios de codificación de un mensaje pueden publicarse sin poner en peligro la seguridad del código. El código de Rivest/Shamir/Adleman está basado en la relativa facilidad de determinar que dos números grandes son primos en comparación con la enorme dificultad práctica de factorizar su producto cuando no se tiene conocimiento previo de cómo se construyó. Si el producto de dos números primos de cien cifras cada uno fuese hecho público, cualquiera podría codificar un mensaje utilizando dicho número de 200 cifras. Pero la decodificación sólo sería posible si fuesen conocidos los dos factores. Existen sistemas criptográficos de clave pública que no dependen de una factorización, pero la seguridad del sistema de Rivest/Shamir/Adleman reside en la intratabilidad del problema de factorización, mientras que su operatividad descansa en la suposición de que los dos factores de 100 cifras son realmente primos. Así pues, las pruebas eficientes de primalidad que no dependen de la factorización tienen un gran valor para el sistema criptográfico.

Todas las pruebas de primalidad que no dependen de factorización se remontan a un teorema enunciado por primera vez por Pierre de Fermat en una carta a su amigo Bernard Frénicle de Bessy con fecha 18 de octubre de 1640. El teorema, usualmente denominado pequeño teorema de Fermat, enuncia que si n es un número primo y b un número entero cualquiera, entonces $b^n - b$ es un múltiplo de n. Por ejemplo, si n es igual a 7 y b a 2, el teorema asegura, correctamente, que $2^7 - 2 = 126$ es un múltiplo de 7.

La importancia del teorema como criterio de primalidad es el enunciado, lógicamente equivalente, que si $b^n - b$ no es múltiplo de n entonces n es compuesto. Para n igual a 4 y b igual a 3, la expresión $3^4 - 3$ es igual a 78, que al dividir por 4 da resto 2; así pues, el pequeño teorema de Fermat permite concluir indirectamente que 4 no es primo.

Aunque el pequeño teorema es un resultado potente y fundamental, existen varias demostraciones elementales del mismo, de las que más adelante expondremos una. El teorema permite enunciar propiedades de números que ni siquiera es posible escribirlos en forma decimal. Por ejemplo, del hecho que $2^{44.497} - 1$ sea primo, el teorema permite concluir que si sustraemos 3 a 3 elevado a $2^{44.497} - 1$, entonces el resultado es divisible por $2^{44.497} - 1$. El número que resulta de efectuar la exponenciación citada adquiere una magnitud tal que ni siquiera puede escribirse en forma decimal. Además, el proceso de división para hallar el cociente no podría ser llevado a cabo por ningún computador físicamente concebible.

Estos números enormes pueden tratarse por medio de la rueda aritmética, inventada por Carl Friedrich Gauss. Este formuló la aritmética modular, en la cual es irrelevante la magnitud de un número y sólo cuenta la parte de la última vuelta de la rueda empleada para llegar a dicho número. Un número n es expresado como el resto de dividir n por un número dado m, al que se llama módulo. Dicho resto se designa por n módulo m, o n (mód m). El número m representa el papel del número de pasos de la rueda, n representa la magnitud absoluta de un número y el resto n (mód m) la fracción de la última vuelta necesaria para alcanzar n.

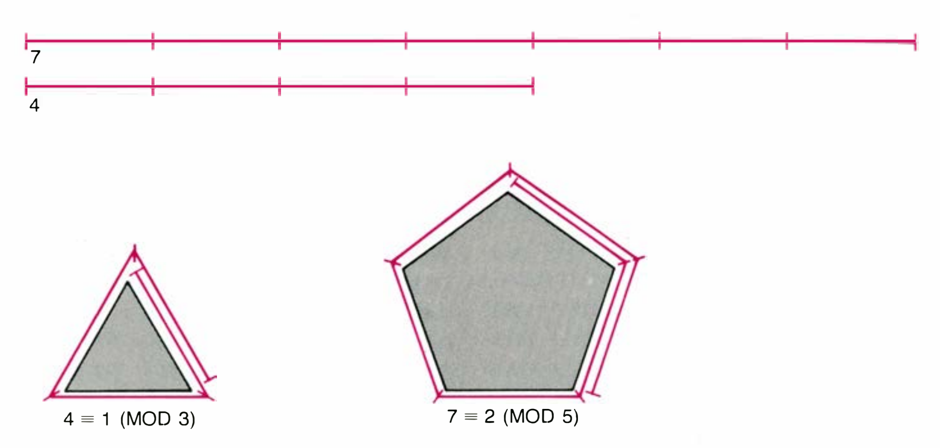
Muchas leyes de la aritmética ordinaria se transfieren a la aritmética modular. En particular, en ésta se puede sumar y multiplicar a condición de expresar los resultados como congruencias. Esto es, todos los números que dan el mismo resto con respecto a un módulo se consideran congruentes con respecto a dicho módulo.

En aritmética ordinaria 6 más 7 es 13. Lo que tiene su análogo en aritmética modular; por ejemplo de módulo 5; así: 6 (mód 5) + 7 (mód 5) es congruente a $1 + 2 = 3$, y 13 también es congruente con 3 (mód 5). De forma parecida, $4 \times 5 = 20$, mientras que en la aritmética de módulo 3 la multiplicación se efectúa de este modo: 4 (mód 3) $\times$ 5 (mód 3) es igual a 1 (mód 3) $\times$ 2 (mód 3), es decir, igual a 2. Por otra parte 20 (mód 3) es también 2.

Con la notación de Gauss el pequeño teorema de Fermat nos dice que si n es primo entonces $b^n - b$ es 0 (mód n). En otras palabras, que $b^n - b$ es un múltiplo de n. La ventaja de la notación de Gauss es que las reglas de la aritmética modular hacen posible calcular el valor de $b^n - b$ módulo n sin tener que dividir $b^n - b$ por n. Para un número tal como $2^7 - 2$ las ventajas del sistema de Gauss no parecen significativas debido a que la división directa es fácil. Mas para hallar el resto de un número tal como $3^{1037} - 3$ cuando se divide por 1037, la aritmética modular se hace casi indispensable.

La esencia del problema es hallar el valor de 3^{1037} (mód 1037). En aritmética modular no es necesario calcular el valor del número 3^{1037}, que es enorme. Basta con aplicar repetidamente el hecho siguiente: el resto del cuadrado de un número es congruente con el cuadrado del resto del número.

Por ejemplo, una vez calculado 3^8 (mód 1037), podemos calcular 3^{16} (mód 1037) hallando el cuadrado del resto de 3^8 y después el resto de este cuadrado (mód 1037). De esta manera se pueden hallar los restos (mód 1037) de 3, 3^2, 3^4, 3^8 y así sucesivamente hasta 3^{1024}. El número 3^{1037} es igual a $3^{(1024+8+4+1)}$, igual, a su vez, a $3^{1024} \times 3^8 \times 3^4 \times 3$,



3. LA ARITMETICA MODULAR es un sistema de cálculo con importantes aplicaciones a las pruebas de primalidad. En aritmética modular la única cosa que interesa de un número *n* es su resto al dividirlo por cierto módulo *m*. La magnitud absoluta del número *n* no se tiene en cuenta. Un ejemplo familiar de aritmética modular lo encontramos en la medición del tiempo, en el que las horas quedan designadas por valores módulo 12. El signo formado por tres rayas horizontales paralelas se lee "es congruente con"; los números a ambos lados del signo dan el mismo resto al ser divididos por el módulo. Por ejemplo, la expresión 4 ≡ 1 (mód 3) significa que 4 y 1 dan el mismo resto cuando se dividen por 3 (en este caso 1).

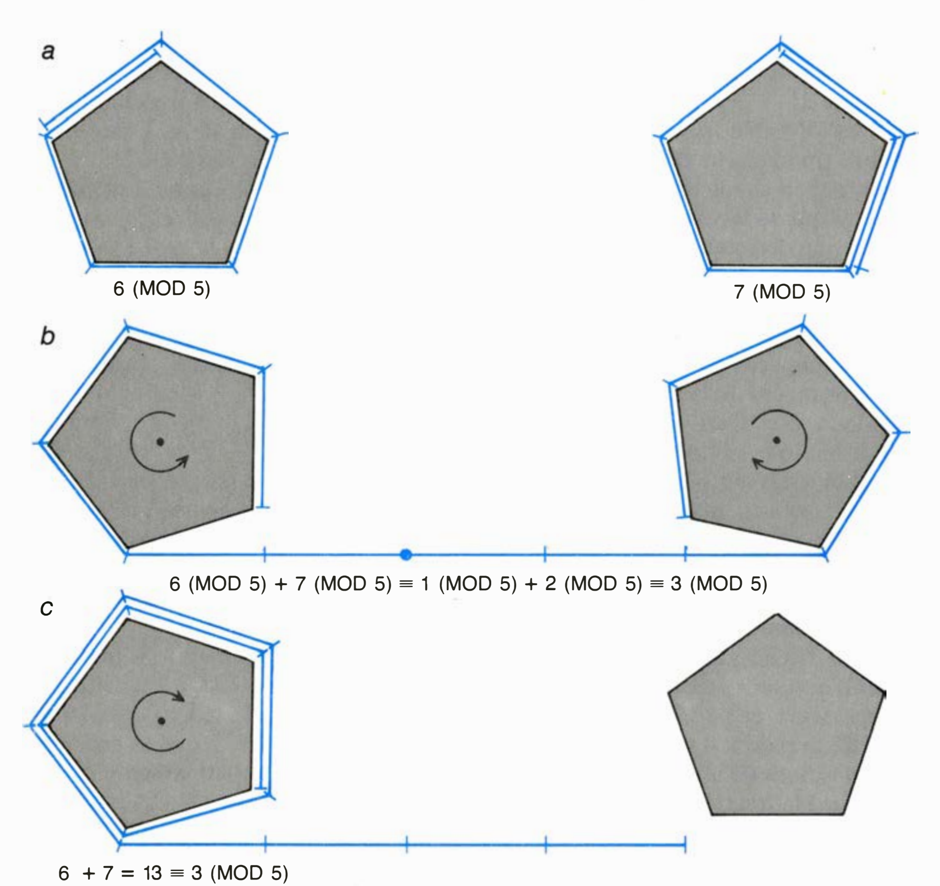
por la ley de los exponentes. Así que 3^{1037} (mód 1037) es congruente con 3^{1024} (mód 1037) × 3^8 (mód 1037) × 3^4 (mód 1037) × 3 (mód 1037). Después de efectuar todas estas operaciones se halla que 3^{1037} es congruente con 845 (mód 1037) y por consiguiente $3^{1037}-3$ es congruente con 842 (mód 1037) [*véase la figura 6*]. Puesto que este resto es no nulo, el pequeño teorema de Fermat nos permite concluir que 1037 es compuesto. Por supuesto que el procedimiento no da ninguna pista para hallar los factores de 1037.

Resulta entretenido servirse de un calculador programable para llevar a cabo este procedimiento. A fin de evitar errores de redondeo conviene limitarse a números n con un máximo de cifras igual a la mitad del número de cifras de la pantalla. Con un computador grande, el cálculo se puede efectuar rápidamente aunque el número tenga miles de cifras. De este modo, la prueba de Fermat permite a menudo decidir que ciertos números, incluso muy grandes, son compuestos.

La demostración del pequeño teorema de Fermat se sigue de una simple consecuencia del teorema fundamental de la aritmética, a saber, que si un número primo divide un producto de varios números, entonces también divide a uno de sus factores. Por ejemplo, 4 × 9 = 36 es divisible por 3, y por supuesto uno de los factores (en este caso 9) es también divisible por 3. El enunciado no es correcto para números compuestos: 4 × 9 es divisible por 6, pero este número no divide ni a 4 ni a 9.

Para demostrar el teorema de Fermat según el cual $b^n - b$ es un múltiplo de n cuando n es primo, obsérvese que $b^n - b$ es igual a $b \times (b^{n-1} - 1)$. De aquí que si b mismo es un múltiplo de n, entonces también lo es $b^n - b$. Así que sólo es necesario demostrar el teorema en el caso que b no sea múltiplo de n, condición que se supondrá en lo que sigue.

La idea básica es que si formamos el producto de los números b, $2b$, $3b$ y así hasta $(n-1)b$, entonces este producto es igual a $b^{n-1}(n-1)!$ Por otra parte, del teorema fundamental de la aritmética resulta que los restos (mód n) de b, $2b$, $3b$, y así sucesivamente hasta $(n-1)b$ son los números 1, 2, 3, y así sucesivamente hasta $n-1$, posiblemente en otro orden. Por consiguiente, $(b^{n-1} - 1)(n-1)!$ es un múltiplo de n. Puesto que el número primo n no divide a ninguno de los números desde 1 hasta $n - 1$, pero divide al producto $(b^{n-1})(n-1)!$, otra aplicación del teorema fundamental de la aritmética permite inferir que n divide a $b^{n-1} - 1$. Como $b^{n-1} - 1$ es un factor de $b^n - b$, el teorema está demostrado.



4. ADICION EN ARITMETICA MODULAR; esta operación se realiza de modo similar a la de la aritmética ordinaria. Se determinan los restos de cada sumando respecto del módulo y se suman los resultados. Si la suma es mayor que el módulo se determina el resto de su suma. En la ilustración se halla la longitud total de dos hilos, uno de longitud 6 y el otro 7. En aritmética módulo 5 no se tienen en cuenta el número de vueltas que el hilo da alrededor del pentágono y sólo se considera relevante la longitud de hilo que sobra tras la última vuelta. Así, módulo 5, 6 es congruente con 1 y 7 lo es con 2 (*a*). Cuando se desenrollan los restos de hilo y se conectan, la longitud total de los dos restos es 3 (*b*). Cuando se arrolla todo el hilo alrededor de un pentágono, la longitud de hilo que sobra, después de haberse agotado todas las vueltas posibles, es 3, de donde la suma ordinaria de 6 y 7 es congruente con 3 módulo 5 (*c*). La suma de los restos de dos números es congruente con el resto de la suma ordinaria de los dos números.

Podría parecer que el teorema de Fermat resuelve completamente el problema de decidir si un número es primo. Pero esto no es así. Si para cierto número b ocurre que $b^n - b$ da resto no nulo cuando lo dividimos por n, entonces n es ciertamente compuesto. Pero si $b^n - b$ es un múltiplo de n, ¿podemos afirmar que n es primo?

Algunos ejemplos podrían sugerir que la respuesta es afirmativa: $2^2 - 2$ es un múltiplo de 2, $2^3 - 2$ es un múltiplo de 3, $2^5 - 2$ es múltiplo de 5, y los números 2, 3 y 5 son primos. Esta pauta fue descubierta hace unos 2500 años por matemáticos chinos, afirmando que si $2^n - 2$ es un múltiplo de n entonces n es primo. Este enunciado fue considerado correcto por Gottfried Wilhelm Leibniz, quien estudió las pautas binarias del *I Ching*. Sin embargo, en 1819 el matemático francés Pierre Frédéric Sarrus señaló que $2^{341} - 2$ es un múltiplo de 341, aun cuando 341 es compuesto (es el producto de 11 y 31). Desde el trabajo de Sarrus se han descubierto muchos otros contraejemplos en los que intervienen valores distintos de la base b: $3^{91} - 3$ es un múltiplo del número compuesto 91 y $4^{15} - 4$ es divisible por 15, que también es compuesto. Todos estos ejemplos se pueden verificar utilizando un pequeño calculador y empleando la aritmética modular descrita más arriba.

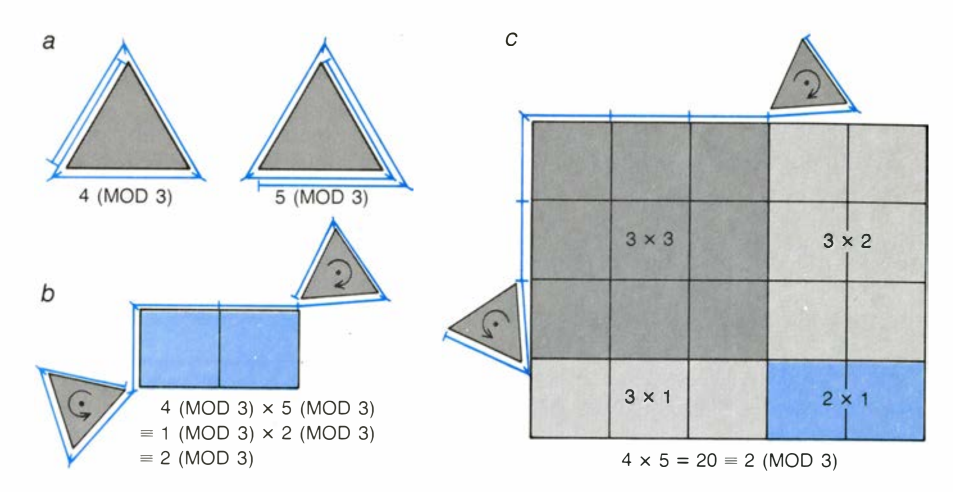
Los números compuestos que la prueba de Fermat no detecta como tales, para un valor dado de b, se denominan seudoprimos respecto de b. El número 341 es seudoprimo respecto de 2, 91 lo es respecto de 3 y 15 respecto de 4. Ocurre que, en relación con una base dada b, existen infinitos seudoprimos. Hay incluso números compuestos, como 561 (igual al producto de 3, 11 y 17) y 1729 (producto de 7, 13 y 19), que son seudoprimos con respecto a cualquier base b. A estos números se les denomina números de Carmichael, en ho-

nor del matemático americano R. D. Carmichael, quien descubrió sus propiedades en 1909.

La existencia de los números de Carmichael pone fin a la esperanza de que la prueba de Fermat, al menos en su concepción original, discrimine entre números primos y compuestos. Sin embargo, los números de Carmichael son extraordinariamente raros; los propios seudoprimos respecto a una base b escasean en comparación con los primos. Jan Bohman, de la Universidad de Lund, ha mostrado que existen 882.206.716 primos menores que 20.000 millones. Pero John L. Selfridge, de la revista *Mathematical Reviews*, y Samuel S. Wagstaff, Jr., de la Universidad de Georgia, han calculado que existen tan sólo 19.865 seudoprimos respecto de 2 que son menores de 20.000 millones. Si la prueba de Fermat fuese llevada a cabo con base 2 para todos los números menores que 20.000 millones, la proporción de errores sería de 1 por millón

La escasez de seudoprimos respecto de 2 entre los números inferiores a 20.000 millones sugiere que un número que pase la prueba de Fermat con base 2 será, probablemente, primo. Además, si se trata de un número compuesto que pasa la prueba con base 2, posiblemente no ocurrirá lo mismo con base 3. Desearíamos poder afirmar que, por aplicación de la prueba de Fermat con base 3, se reduce significativamente la probabilidad de que un número compuesto que ha pasado la prueba con base 2, no sea primo. Sin embargo, debido a que dichas pruebas no son necesariamente independientes, la prueba de Fermat con base 3 no descarta, quizá, números compuestos que tampoco habían sido descartados por la prueba con base 2.

D. H. Lehmer, de la Universidad de California en Berkeley, e independientemente Robert M. Solovay, del Instituto de Tecnología de California, y Volker Strassen, del Instituto Federal de Tecnología de Suiza, acaban de desarrollar una variante de la prueba de Fermat que satisface el requerimiento de que las pruebas correspondientes a distintas bases sean independientes. Una de las características de la prueba es que si el número n es compuesto, entonces aparecerá como compuesto al menos para la mitad de los valores de b seleccionados entre 1 y n. Así pues, si se seleccionan, por ejemplo, 100 bases distintas al azar y con cada una de ellas llevamos a cabo la prueba de Lehmer/Solovay/Strassen, entonces la probabilidad de que un número compuesto n seleccionado al azar pase las 100 pruebas es menor que 1 entre 2^{100}, aproximadamente 1 entre 10^{30}.



5. TAMBIEN LA MULTIPLICACION en aritmética modular se realiza de modo similar a la multiplicación ordinaria. Aquí se multiplican 4 y 5 módulo 3, que se representan por hilos arrollados alrededor de un triángulo. El resto de cada número es la longitud de hilo que sobra después de agotar todas las vueltas completas posibles (*a*). El resto de 4, módulo 3, es 1 y el de 5 es 2. El producto de los dos restos es el área del rectángulo cuyos lados son iguales a dichos restos. En nuestro caso 2 × 1 = 2 (*b*). Por otra parte, el producto de 4 y 5 es el área del rectángulo cuyos lados miden 4 y 5. El resto de este producto (*c*) se obtiene despreciando el área de cualquier rectángulo menor, dentro del rectángulo, uno de cuyos lados (al menos) es igual a la longitud de hilo necesario para dar una vuelta alrededor del triángulo (*regiones en gris*). El área del rectángulo restante (*región de color*) es el resto del producto de 4 y 5 módulo 3. Afirma la regla general que el producto de los restos de dos números es congruente con el resto de su producto ordinario.

Solovay y Strassen consideraron que su prueba era una "prueba de primalidad de Monte Carlo". (Ciertos métodos probabilísticos en matemáticas y física se denominan así por la ciudad conocida por los juegos de azar.) Desde un punto de vista práctico, tal asignación se diría correcta; por ejemplo, en las aplicaciones criptográficas podría parecer que el funcionamiento de un código no se verá disminuido por la probabilidad despreciable de que los números en los que se basa no sean, a la postre, primos. Pero también se ha argumentado, basándose en premisas filosóficas que considero incorrectas, que habida cuenta de que toda demostración matemática es susceptible a errores, y por tanto corregible, debiera aceptarse como una prueba matemática una fuerte verificación, en términos probabilísticos, de la primalidad de un número.

Hay razones para creer que la probabilidad de que argumentos aceptados como demostraciones resulten erróneos supera de lejos la proporción de 1 entre 10^{30}. La historia de la matemática ofrece múltiples ejemplos de "demostraciones" cuya falsedad o error se descubrió más tarde. Entre verificaciones probabilísticas y demostraciones matemáticas media una diferencia cualitativa, de máxima importancia para los matemáticos. Una demostración es un argumento deductivo en el que cada paso es una consecuencia lógica de pasos precedentes. La importancia de la demostración yace no sólo en que se pueda ver que la conclusión es válida, sino también en que su conclusión válida se derivará necesariamente del argumento. Pienso que lo que sugiere la idea de una prueba de primalidad de Monte Carlo es que el concepto de demostración y el de certeza son muy distintos entre sí.

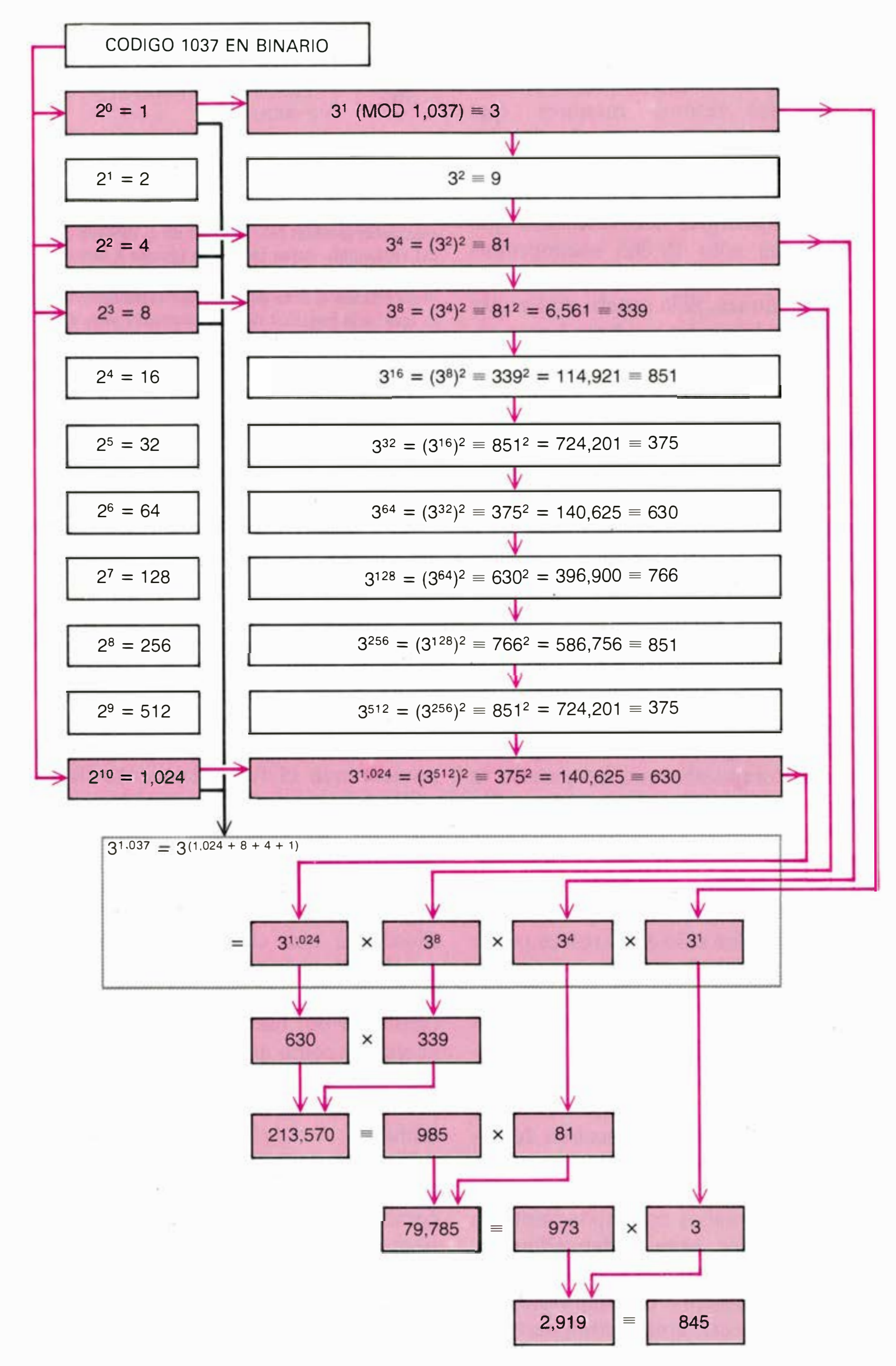
En 1876 el matemático francés Edouard A. Lucas halló una prueba de primalidad rigurosa para todo número n. Supongamos que existe un número b tal que b^{n-1} es congruente con 1 (mód n), pero de modo que $b^{(n-1)/p}$ no es congruente con 1 (mód n) para todo factor primo p de $n - 1$. Entonces, como demostró Lucas, n es primo.

Supongamos por ejemplo que el número en cuestión es 257. Entonces $n - 1$ es $256 = 2^8$, de modo que 2 es el único divisor primo de $n - 1$. A fin de probar que n es primo tenemos que hallar un número b tal que b^{n-1} es congruente con 1 (mód 257) y que $b^{256/2}$ no es congruente con 1 (mód 257). Aunque la prueba de Lucas no indica cómo hallar un número b específico, dado el número n siempre hay muchos números b que satisfacen las condiciones del teorema de Lucas. Casi siempre se hallará un tal b por medio de una búsqueda al azar. Cuando b es 3, por ejemplo, 3^{256} es congruente con 1 (mód 257), pero $3^{256/2}$ es congruente con 256 (mód 257). Por tanto, 257 será primo. Aunque la prueba de Lucas también es una prueba de Monte Carlo, en el sentido

de que b se selecciona al azar, proporciona una prueba rigurosa una vez hallado un número b.

Hay un aspecto de la prueba de Lucas que la convierte en inaplicable a números que no tengan una forma especial: su aplicación exige conocer todos los factores primos p de $n - 1$. Por supuesto que si sospechamos que n es primo entonces n será impar y por tanto $n - 1$ será divisible por 2. Raramente basta este punto de arranque; la prueba de Lucas suele resultar inviable a no ser que $n - 1$ pueda factorizarse con facilidad, como en el ejemplo anterior.

Si todos los factores primos de $n + 1$ pueden ser hallados sin la dificultad que rodea a los de $n - 1$, se determinará la primalidad de n por otra prueba, también propuesta por Lucas. Prueba

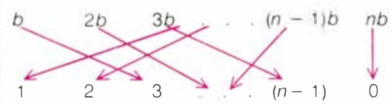


6. APLICACION DE LA ARITMETICA MODULAR para evitar cálculos que de otro modo serían, además de muy largos, susceptibles de errores. El diagrama muestra cómo se puede calcular el resto de 10^{1037}, al dividirlo por 1037, sin calcular 10^{1037}. El método se basa en el hecho que el resto del cuadrado de un número es congruente con el cuadrado del resto. De este modo se pueden hallar los residuos, módulo 1037, de 3, 3^2, 3^4, 3^8, 3^{16}, y así sucesivamente hasta 3^{1024}, sin más que calcular repetidamente el resto del cuadrado del resto de una potencia de 3. Puesto que 3^{1037} es igual a $3^{1024} \times 3^8 \times 3^4 \times 3$, el resto de 3^{1037} (mód 1037) es igual al resto del producto de los restos de 3^{1024}, 3^8, 3^4 y 3, módulo 1037. Este procedimiento se puede llevar a cabo con un calculador programable. (Ilustración de Ian Warpole.)

que mejoraría, en 1930, Lehmer. Para números a los que se puede aplicar, esta prueba de Lucas/Lehmer la efectúa con extraordinaria rapidez un computador grande. Los números primos mayores que se conocen, de la forma $2^p - 1$, donde p es a su vez primo, se encontraron merced a dicha prueba. Se trata de los números de Mersenne, en honor de Marin Mersenne, un matemático francés del siglo XVII, quien dio una lista de números primos p para los cuales $2^p - 1$ es primo. Está claro que si n es un número de Mersenne, entonces $n + 1$ sólo posee el factor primo 2.

En 1975, John Brillhart, de la Universidad de Arizona, Lehmer y Selfridge construyeron una prueba de primalidad para el número n que sólo depende del conocimiento de algunos factores primos de $n - 1$ o de $n + 1$. Hugh C. Williams, de la Universidad de Manitoba, ha elevado esta clase de pruebas a la categoría de arte: para verificar la primalidad de n basta ahora conocer factorizaciones parciales de $n^2 + 1$, $n^2 - n + 1$ o $n^2 + n + 1$. Pero si ninguno de estos factoriza con facilidad, dichas pruebas resultan ineficientes. Aunque a muchos números de 100 cifras se les puede aplicar este tipo de pruebas, se estima que para ciertos números primos de 100 cifras el tiempo de computación requerido tardaría un siglo. Se hace patente la necesidad de una prueba de primalidad que no dependa de la forma especial del número en cuestión.

CONSIDEREMOS LOS NUMEROS b, $2b$, $3b$, ..., $(n-1)b$, nb
SUS RESTOS (MOD n) SON LOS COEFICIENTES DE b EN OTRO ORDEN.



ASI QUE $b \times 2b \times ... \times (n-1)\,b$ (MOD n) $\equiv 1 \times 2 \times ... \times (n-1)$ (MOD n).
POR CONSIGUIENTE $[b \times 2b \times ... \times (n-1)b] - [1 \times 2 \times ... \times (n-1)] \equiv 0$ (MOD n).
PERO $b \times 2b \times ... \times (n-1)b = b^{n-1}(n-1)!$ y $1 \times 2 \times ... \times (n-1) = (n-1)!$
POR TANTO $[b \times 2b \times ... \times (n-1)b] - [1 \times 2 \times ... \times (n-1)] =$
$[b^{n-1}(n-1)!] - [(n-1)!] = (b^{n-1} - 1)(n-1)! \equiv 0$ (MOD n).
ENTONCES $(b^{n-1} - 1)(n-1)!$ ES UN MULTIPLO DE n, LO QUE ES SUFICIENTE PARA DEMOSTRAR EL TEOREMA DE FERMAT.

7. DEMOSTRACION DEL "PEQUEÑO TEOREMA" DE FERMAT. Se sigue de la aplicación del teorema fundamental de la aritmética y de las reglas de multiplicación en aritmética modular. El pequeño teorema de Fermat afirma que si n es un número primo y b un número entero cualquiera, entonces $b^n - b$ es un múltiplo de n. Si b es un múltiplo de n la conclusión del teorema es inmediata, ya que $b^n - b$ es igual a $b(b^{n-1} - 1)$, b es un factor de $b^n - b$ y por tanto $b^n - b$ es un múltiplo de n. Si b no es un múltiplo de n, entonces es suficiente demostrar que $b^{n-1} - 1$ es múltiplo de n. Un corolario del teorema fundamental de la aritmética afirma que si un número primo divide un producto de varios números, entonces dicho número también divide a uno de los factores. En consecuencia, n no divide a $(n-1)!$, ya que no divide a ninguno de los números desde 1 hasta $n - 1$. Así pues $b^n - b$ será divisible por n, y por tanto también $b^{n-1} - 1$, si podemos demostrar que $(b^{n-1} - 1)(n-1)!$ es divisible por n. A tal fin consideremos los números b, $2b$, $3b$, y así hasta nb. Dos números distintos de esta sucesión, digamos ib y jb, no pueden dar el mismo resto al dividirlos por n, ya que, si tal ocurriera, ib-jb, que es igual a $(i$-$j)b$, sería un múltiplo de n, de donde, por el teorema fundamental de la aritmética, resultaría que i-j sería divisible por n, lo cual es imposible. Teniendo en cuenta que el resto de nb, al dividir por n, es 0, resulta que los restos de los números b, $2b$, ..., $(n-1)b$ tienen que ser los números entre 1 y $n - 1$, posiblemente en otro orden. El diagrama muestra cómo la operación de dividir los números b, $2b$, ..., $(n-1)b$ por n, y después tomar el resto, establece una correspondencia entre los números b, $2b$, ..., $(n-1)b$ y los números 1, 2, ..., $(n-1)$, posiblemente permutados. Como los restos de los números en cada conjunto son los mismos salvo el orden, el resto del producto $b \times 2b \times ... \times (n-1)b$ es igual al resto del producto $1 \times 2 \times ... \times (n-1)$. La diferencia de los dos productos es múltiplo de n, pues los restos se cancelan. Pero la diferencia es igual a $(b^{n-1} - 1)(n-1)!$, según muestra una sencilla transformación algebraica. Por tanto, $b^{n-1} - 1$ es divisible por n, lo que termina la demostración del pequeño teorema de Fermat para los b enteros.

342 ES SEUDOPRIMO RESPECTO DE 2
$2^{341} = 2^{256} \times 2^{64} \times 2^{16} \times 2^4 \times 2$
$\equiv 64 \times 16 \times 64 \times 16 \times 2$ (MOD 341)
$\equiv 2$ (MOD 341)
DE AQUI QUE $2^{341} - 2 \equiv 0$ (MOD 341)
POR CONSIGUIENTE 341 PASA LA PRUEBA DE FERMAT DE BASE 2
PERO $341 = 11 \times 31$

561 ES SEUDOPRIMO RESPECTO DE CUALQUIER BASE
$2^{561} = 2^{512} \times 2^{32} \times 2^{16} \times 2$
$\equiv 103 \times 103 \times 460 \times 2$ (MOD 561)
$\equiv 2$ (MOD 561)
ASI QUE $2^{561} - 2 \equiv 0$ (MOD 561)
POR TANTO 561 PASA LA PRUEBA DE FERMAT DE BASE 2
PERO $561 = 3 \times 11 \times 17$
$3^{561} = 3^{512} \times 3^{32} \times 3^{16} \times 3$
$\equiv 273 \times 273 \times 69 \times 3$ (MOD 561)
$\equiv 3$ (MOD 561)
ASI QUE $3^{561} \times 3 \equiv 0$ (MOD 561)
EN CONSECUENCIA 561 PASA LA PRUEBA DE FERMAT DE BASE 3

8. DE UN NUMERO DECIMOS QUE ES SEUDOPRIMO cuando pasa la prueba de primalidad derivada del pequeño teorema de Fermat, para algún valor de b, a pesar de ser en realidad compuesto. Un seudoprimo respecto de una base b es un número compuesto n que divide a $b^n - b$. Pierre Frédéric Sarrus fue el primero en señalar que 341, que es el producto de 11 y 32, es seudoprimo respecto de la base 2. Aquí se utiliza la aritmética modular para verificar que $2^{341} - 2$ es divisible por 341; el procedimiento es similar al utilizado en la figura 6. Los números de Carmichael, llamados así en honor del matemático R. D. Carmichael, son seudoprimos respecto de cualquier base b. El número 561, que es el producto de 3, 11 y 17, es el menor número de Carmichael; aquí se muestra que es seudoprimo respecto de 2 y 3.

Adleman y Robert S. Rumely, de la Universidad de Georgia, desarrollaron en 1980 una prueba que ha mejorado radicalmente la eficiencia de las pruebas de primalidad de números grandes y carentes de forma especial. Tal como se formuló inicialmente, la prueba era probablemente capaz de decidir la primalidad de cualquier número con un número de cifras entre 50 y 100 en un tiempo comprendido entre 4 y 12 horas (con un computador grande). Henri Cohen, de la Universidad de Burdeos, y Hendrik W. Lenstra, Jr., de la Universidad de Amsterdam, han perfeccionado desde entonces la prueba en varios puntos; ahora se puede llevar a cabo 1000 veces más aprisa. Con un computador Cyber 170-750 (de Control Data Corporation), un número de 100 cifras se decide en 40 segundos.

¿Cómo consigue la prueba de Adleman/Rumely semejante rendimiento? Para llegar hasta sus detalles se impone dominar la teoría de los números algebraicos, pero en su esencia es bastante similar a la prueba de Fermat. Se construyen dos números auxiliares, un número I llamado número inicial y un número E llamado número euclídeo. El número I es un producto de varios números primos; por ejemplo: $2 \times 3 \times 5 \times 7 = 210$. El número E se llama euclídeo porque su construcción recuerda la demostración de Euclides de la existencia de infinitos números primos. E es el producto de todos los números primos p, q, r, y así sucesivamente, tales que $p - 1$, $q - 1$, $r - 1$, etcétera, son todos los factores de I. Por ejemplo, 70 es un factor de 210, y como $70 = 71 - 1$ y

71 es primo, 71 es un factor E. Los divisores de 210 iguales a un número primo disminuido en una unidad son 1, 2, 6, 10, 30, 42, 70 y 210. Por consiguiente, E es el producto de los números primos 2, 3, 7, 11, 31, 43, 71 y 211, es decir, 9.225.988.926. El número E se ha de construir de suerte que sea mayor que la raíz cuadrada del número n en cuestión. En el ejemplo anterior, con número inicial 210, el método de Adleman/Rumely sería válido para números n no superiores a 10^{19}. Por ser el tiempo de computación proporcional a una potencia de I, el número inicial debe seleccionarse tan pequeño cuanto sea posible.

Asistimos pues a una especie de tensión dinámica entre los números I y E. Para que la prueba sea válida, E ha de ser grande; para que sea rápida, I ha de ser pequeño. Además, como E depende de I, su selección no puede ser independiente. El número 210 constituye un buen ejemplo de número inicial I, ya que se trata de un número bastante pequeño que posee muchos factores que son un número primo disminuido en una unidad. Para demostrar que la prueba de Adleman/Rumely es siempre rápida hubo que verificar que siempre se pueden hallar números E e I, y estimar su magnitud.

Por coincidencia ya se habían hecho ciertos trabajos sobre esta cuestión. En 1955 Karl Prachar, de la Universidad Agrícola de Viena, había mostrado que existen infinitos números que poseen un gran número de factores que son números primos disminuidos en una unidad. A fin de aplicar el resultado de Prachar a la prueba original de Adleman/Rumely fue necesario mostrar que los números I se podían construir de modo que no tuviesen factores cuadrados perfectos mayores que 1. Cohen y Lenstra han demostrado recientemente que en su variante de la prueba podían prescindir de esta condición. El resultado de Prachar se ha reforzado con ulteriores hallazgos de Patrick X. Gallagher, de la Universidad de Columbia, y Enrico Bombieri, de Princeton. Andrew M. Odlyzko, de los laboratorios Bell, y el autor han analizado la construcción de los números susceptibles de aparecer como I en una prueba de primalidad.

Después de la construcción de los números I y E, se aplican ciertas pruebas, similares a las de Fermat, a pares de números primos p y q, tales que p es un divisor de E y q un divisor de $p - 1$. La prueba no se lleva a cabo con números enteros, sino con números enteros algebraicos. Un entero algebraico es un número complejo que es la raíz de un polinomio de coeficientes enteros cuyo coeficiente de grado máximo es 1. Por ejemplo, $\sqrt{2}$, i (una raíz cuadrada de -1) y $(-1 + i\sqrt{3})/2$ son enteros algebraicos, ya que satisfacen las ecuaciones algebraicas $x^2 - 2 = 0$, $x^2 + 1 = 0$ y $x^3 - 1 = 0$, respectivamente.

Si un número n no pasa la prueba correspondiente a uno de los pares de primos p y q, entonces n es compuesto. Si n pasa todas las pruebas, no queda decidido todavía si n es primo, pero el número de posibles factores que se requiere comprobar es pequeño. Adleman y Rumely han demostrado que todo número compuesto n que pase todas las pruebas necesariamente tiene factores primos en un conjunto de I elementos. Lenstra, por su parte, ha probado que los números en el conjunto son los

PRIMOS DE MERSENNE INFERIORES A $2^{62.982}$

VALORES DE p TALES QUE $2^p - 1$ ES PRIMO	$2^p - 1$	CUANDO SE DEMOSTRO QUE LO ERA	AUTORES	MAQUINA USADA
2	3	ANTIGUEDAD	MENCIONADO EN LOS ELEMENTOS DE EUCLIDES	
3	7			
5	31			
7	127			
13	8191	1461	MENCIONADO EN EL CODICE LAT. MONAC. 14908	
17	131.071	1588	PIETRO ANTONIO CATALDI	
19	524.287			
31	2.147.483.647	1772	LEONHARD EULER	
61	19 CIFRAS	1883	I. M. PERVOUCHINE	
89	27 CIFRAS	1911	R. E. POWERS	
107	33 CIFRAS	1914	R. E. POWERS, E. FAUQUEMBERGE	
127	39 CIFRAS	1876-1914	ÉDOUARD LUCAS, E. FAUQUEMBERGE	
521	157 CIFRAS	1952	RAPHAEL M. ROBINSON	SWAC
607	183 CIFRAS			
1279	386 CIFRAS			
2203	664 CIFRAS			
2281	687 CIFRAS			
3217	969 CIFRAS	1957	HANS RIESEL	BESK
4253	1281 CIFRAS	1961	ALEXANDER HURWITZ	IBM-7090
4423	1332 CIFRAS			
9689	2917 CIFRAS	1963	DONALD B. GILLIES	ILLIAC-II
9941	2993 CIFRAS			
11.213	3376 CIFRAS			
19.937	6002 CIFRAS	1971	BRYANT TUCKERMAN	IBM 360/91
21.701	6533 CIFRAS	1978	LAURA NICKEL, CURT NOLL	CDC-CYBER-174
23.209	6987 CIFRAS	1979	CURT NOLL	CDC-CYBER-174
44.497	13.395 CIFRAS	1979	HARRY L. NELSON, DAVID SLOWINSKI	CRAY-1

9. POR NUMERO PRIMO DE MERSENNE se entiende el número primo que es una potencia de 2 disminuida en una unidad. Recibe ese nombre en honor de Marin Mersenne. Han presentado interés desde la antigüedad ya que, como mostró Euclides, si el número $2^p - 1$ es primo, entonces $2^{p-1}(2^p - 1)$ es perfecto, vale decir, constituye la suma de todos sus divisores con excepción del mismo número. Si $2^p - 1$ es primo, entonces p también será primo, pero lo recíproco no es necesariamente cierto. Se desconoce si existen infinitos números primos de Mersenne y si hay infinitos números compuestos de Mersenne. Cabe la probabilidad de que ambos enunciados resulten ciertos. En los últimos años, el crecimiento de la tabla de números primos de Mersenne conocidos ha ido paralelo al crecimiento de la capacidad de computación. De acuerdo con David Slowinski decidir que $2^{8191} - 1$ no es primo llevó 100 horas de computación con el ILLIAC-I, 5,2 horas en el IBM 7090, 49 minutos con el ILLIAC-II, 3,1 minutos con el IBM 360/91 y 10 segundos con el CRAY-1. Slowinski y Harry L. Nelson, del Lawrence Livermore, han examinado los números de Mersenne para todos los valores de p hasta 50.000, sin que hayan identificado entre ellos números primos mayores que $2^{44.497} - 1$. Guy M. Haworth, Steven M. Holmes, David J. Hunt, Thomas W. Lake y Stewart F. Reddaway, de International Computing Limited, acaban de reanudar la búsqueda con ayuda del supercomputador ICL-DAP, que posee 4096 procesadores que operan en paralelo. Hasta ahora, han buscado primos de Mersenne para valores de p hasta 62.982, pero no han encontrado todavía ninguno.

MAGNITUD DEL NUMERO					
PRUEBA DE PRIMALIDAD	20 CIFRAS	50 CIFRAS	100 CIFRAS	200 CIFRAS	1000 CIFRAS
POR DIVISION	2 HORAS	10^{11} AÑOS	10^{36} AÑOS	10^{86} AÑOS	10^{486} AÑOS
LUCAS, BRILLHART-LEHMER-SELFRIDGE, WILLIAMS	5 SEGUNDOS	10 HORAS	100 AÑOS	10^{9} AÑOS	10^{44} AÑOS
ADLEMAN-RUMELY, COHEN-LENSTRA	10 SEGUNDOS	15 SEGUNDOS	40 SEGUNDOS	10 MINUTOS	1 SEMANA

10. TIEMPO REQUERIDO para realizar una prueba de primalidad, ampliamente variable según el tipo de la misma. Aquí se supone que las pruebas las realiza un computador rápido; en particular, en el caso de la prueba por divisiones se admite que el computador efectúa un millón de divisiones por segundo con independencia de la magnitud del número. Los tiempos asignados a las pruebas de tipo similar a la inventada por el matemático francés Edouard A. Lucas corresponden a casos extremos; para números primos de formas especiales, el tiempo suele acortarse más. La mayoría de los tiempos alistados son estimaciones; en las casillas coloreadas figuran tiempos reales experimentados en computador. En la práctica, los tres tipos se combinan para formar una prueba única, algo más rápida que la tercera.

restos (mód E) de las potencias n, n^2, n^3, ..., n^I. Si uno de estos números, que no sea 1 o n, divide a n, entonces n será compuesto; si no, n será primo. Aunque este último paso puede dar la impresión de que el método de Adleman-/Rumely es un método de factorización, debe hacerse hincapié en que la conclusión sólo es válida si n ha pasado todas las pruebas anteriores del tipo de la de Fermat. Muchos números compuestos, quizá todos, no superarán una de estas pruebas y por tanto no poseerán necesariamente un factor que se detectaría en el último paso.

La velocidad y la aplicabilidad completamente general de las nuevas pruebas de primalidad han abierto el camino para una investigación teórica de números hasta ahora inaccesibles incluso a los computadores más rápidos. Supongamos, sin embargo, que números arbitrarios con un número de cifras mucho mayor que 100 son sometidos a las pruebas. ¿Cuánto tiempo tardarán en decidir la primalidad de dichos números? Esta y otras cuestiones similares revisten un evidente interés teórico en la rama de las ciencias de la computación conocida por teoría de la complejidad. De acuerdo con una definición corrientemente aceptada en la teoría de la complejidad, una prueba de primalidad es lenta cuando no se ejecute en lo que se llama tiempo polinómico. Esto es, la prueba se considera lenta a no ser que el tiempo requerido para decidir la primalidad de un número n sea menor que una potencia fija del número de cifras de n. El número de cifras de n se denota $d(n)$. Obsérvese que $d(n)$ es, a su vez, un número y, por tanto, su número de cifras se escribirá $d(d(n))$.

Las pruebas de Adleman/Rumely y de Cohen/Lenstra resultan lentas. El tiempo de computación está acotado por $d(n)$ elevado a $d(d(d(n)))$ multiplicado por una cierta constante c. La expresión $d(d(d(n)))$ es el número de cifras del número de cifras del número de cifras de n. Sea cual sea el valor de c, el exponente superará a cualquier constante prefijada al crecer n indefinidamente; por consiguiente, la cota del tiempo de computación no será polinómica.

Ahora bien, para números relativamente "pequeños", el criterio de la teoría de la complejidad puede inducir a engaño, ya que $d(d(d(n)))$ tiene una tasa de crecimiento extraordinariamente lenta. Incluso para el número 10^{1000}, $d(d(d(n))) = 1$. El primer número n para el cual la expresión es igual a 2 es $10^{999.999.999}$. En otras palabras, para números inferiores a $10^{999.999.999}$, el tiempo de computación está acotado por $d(n)$ elevado a c; por tanto, el tiempo de computación de las nuevas pruebas, aunque no polinómico, "casi" lo es.

Ahora que para números moderadamente grandes las pruebas de primalidad pueden realizarse con rapidez, quizá valga la pena dirigir la atención hacia el problema relacionado de factorización. Todo progreso en este problema tendría implicaciones inmediatas para los sistemas criptográficos basados en ella. Aunque el avance en las pruebas de primalidad no tiene una aplicación directa al problema de factorización, jamás se ha demostrado que ésta sea intratable. No hay ninguna garantía de que alguien no invente mañana un método revolucionario de factorización. Por tanto, una decisión acerca de la seguridad a largo plazo de los sistemas de clave pública fundados en la dificultad de la factorización parte de un juicio subjetivo sobre si habrá o no un progreso importante en los métodos de factorización. Los desarrollos recientes en las pruebas de primalidad sirven para acentuar la potencial vulnerabilidad de tales códigos debida a los avances teóricos.

Las lucernas de Cosa

La evolución de las lucernas, la principal iluminación interior de la civilización clásica, se puede rastrear en un yacimiento arqueológico italiano a lo largo de siete siglos. En esos artefactos hallamos las claves de la cultura de su tiempo

Cleo Rickman Fitch

El empleo por el hombre de iluminación artificial, distinta de la obtenida por el fuego, puede rastrearse hasta los tiempos paleolíticos. Así, el descubrimiento de lucernas de mecha en la cueva francesa de Lascaux. Pero las lucernas no llegaron a usarse en gran escala hasta que el desarrollo de la agricultura hizo asequible algunos productos: aceite de sésamo, aceite de semilla de lino y aceite de oliva. En época de la Grecia clásica, hacia comienzos del siglo v a.C., decenas de miles de estas ingeniosas fuentes de luz, generalmente hechas de "terracota", eran producidas por los artesanos para su distribución regional por Asia Menor, Grecia, Sicilia, zonas de Italia y otros lugares del mundo antiguo. En el área del Mediterráneo, la práctica de fabricar lucernas una a una en la rueda de alfarero fue sustituida por la producción en masa sirviéndose de moldes, ya bastante antes del comienzo de la era cristiana.

La evolución de la forma, función y decoración de estos humildes objetos es interesante tanto en sí misma como por su utilidad para el estudio de la tecnología y la sociedad antiguas. Ahora sabemos cómo las lucernas evolucionaron en Italia central a lo largo de un período de casi 800 años, desde los tiempos republicanos hasta la caída del Imperio Romano, gracias al resultado de una intensa campaña arqueológica llevada a cabo en un yacimiento de la costa del Tirreno. Allí, en un promontorio al otro lado de la bahía, desde un hito costero, el monte Argentario, a unos 120 kilómetros al norte de Roma, fue fundada la *colonia* de Cosa en el año 273 a.C. por los romanos y sus aliados de la Liga Latina después de su victoria de unos años antes sobre dos ciudades-estado etruscas, Vulci y Volsinii. Cosa debía ser un fuerte fronterizo, defensor del territorio conquistado frente a otros enemigos etruscos situados más al norte.

Aunque la colonia fue poblada por unas 2500 familias, no tenía sitio más que para 300 casas, aparte de los habituales edificios públicos, dentro de las grandes murallas de piedra que daban fe del propósito militar de la ciudad. El trazado urbanístico se basaba en manzanas rectangulares con las calles principales discurriendo desde el noroeste hacia el sureste. Se construyó un almacén municipal en el interior, exactamente junto a la puerta noroeste. Otros edificios públicos eran el templo principal de la ciudad (dedicado a Júpiter, Juno y Minerva), otro más pequeño (dedicado a Mater Matuta, diosa del alba), el forum (más o menos en el centro de la ciudad) y una zona de tiendas y baños. Como el sustrato rocoso de Cosa es de piedra caliza, no había fuentes en la ciudad; los habitantes dependían del agua de lluvia recogida en grandes cisternas públicas y en otras particulares más pequeñas habilitadas en cada casa.

La Academia Americana de Roma ha llevado a cabo una campaña de excavaciones en Cosa desde 1948, utilizando el yacimiento como campo de experimentación para formar a arqueólogos clásicos. Un total de 14 campañas de trabajo acaban de concluir, gran parte de ellas bajo la dirección de Frank E. Brown, quien procedía de la Universidad de Yale y desde 1965 a 1969 dirigió la Academia. Las excavaciones hicieron posible la recuperación del material que examinaremos aquí: los fragmentos de miles de lucernas y unas pocas enteras. Algunas se fabricaron en el mismo lugar o se importaron del sur de Italia, Sicilia y norte de Africa. La mayoría, en cambio, fueron fabricadas por alfareros de Roma o de sus alrededores.

Cuando caía la noche en Cosa, como en otras ciudades y aldeas de la época, las casas de los pudientes se alumbraban con candelabros y lucernas hechas de bronce e incluso de metales preciosos, los altares de los templos brillaban con velas votivas y ardían lucernas en las tiendas de vino y tabernas. También había lámparas en las casas corrientes, pero sólo se encendían en ocasiones especiales. Como los campesinos de todas partes, el cabeza de familia gastaba poco en aceite, normalmente yendo a la cama cuando oscurecía y levantándose antes del alba. Aún así había pocos hogares que no tuvieran una lámpara o dos de cerámica. Ahora bien, este material era frágil, por lo cual no resulta sorpresa alguna que la mayoría de las encontradas en Cosa estén rotas.

Pese a toda su sencillez la lucerna fue un invento extraordinario. Los primeros ejemplares eran cuencos poco profundos de barro cocido, hechos en una rueda de alfarero, o a mano. El cuenco se llenaba de aceite hasta la mitad; la mecha flotaba en el líquido. Después, algún alfarero imaginativo apretó el borde del cuenco antes de cocer el barro, formando un pequeño labio, donde un extremo de la mecha podía descansar encima del nivel del aceite. La siguiente innovación consistió en volver el borde del cuenco hacia adentro, menos la parte con el labio, con lo cual se redujo el riesgo de derramar el aceite. Más o menos al mismo tiempo se hizo costumbre fijar un asa a la lámpara en el lado opuesto a donde estaba el labio. El proceso continuó hasta que solamente llegó a haber una pequeña abertura en la parte superior: el orificio para rellenar.

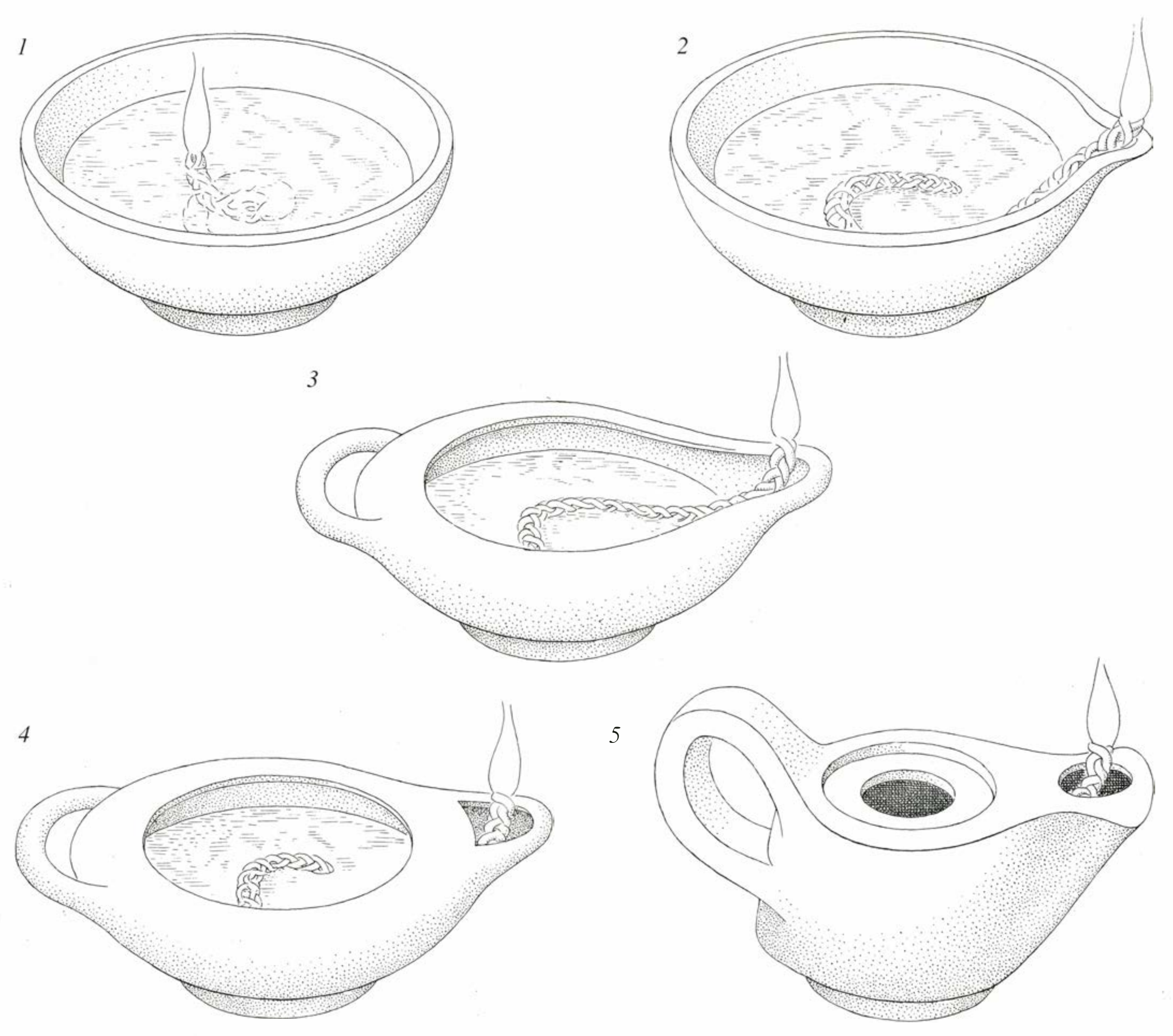
A continuación desapareció el labio y un cono de barro moldeado a mano ocupó su lugar. El alfarero fijaba el cono a un lado de la lámpara mientras todavía estaba blanda y entonces hacía un agujero a través del cono, desde su punta hacia abajo hasta el depósito de aceite de la lámpara, proporcionando así a la vez una boquilla para la mecha. Más tarde, en el siglo III a.C., alfareros de Asia Menor comenzaron la produc-

ción en masa de lucernas moldeadas en dos partes. Como ciertos pasos del proceso de moldeo no necesitan trabajo especializado, la nueva técnica de fabricación pronto fue adoptada en otros lugares. El proceso es el siguiente.

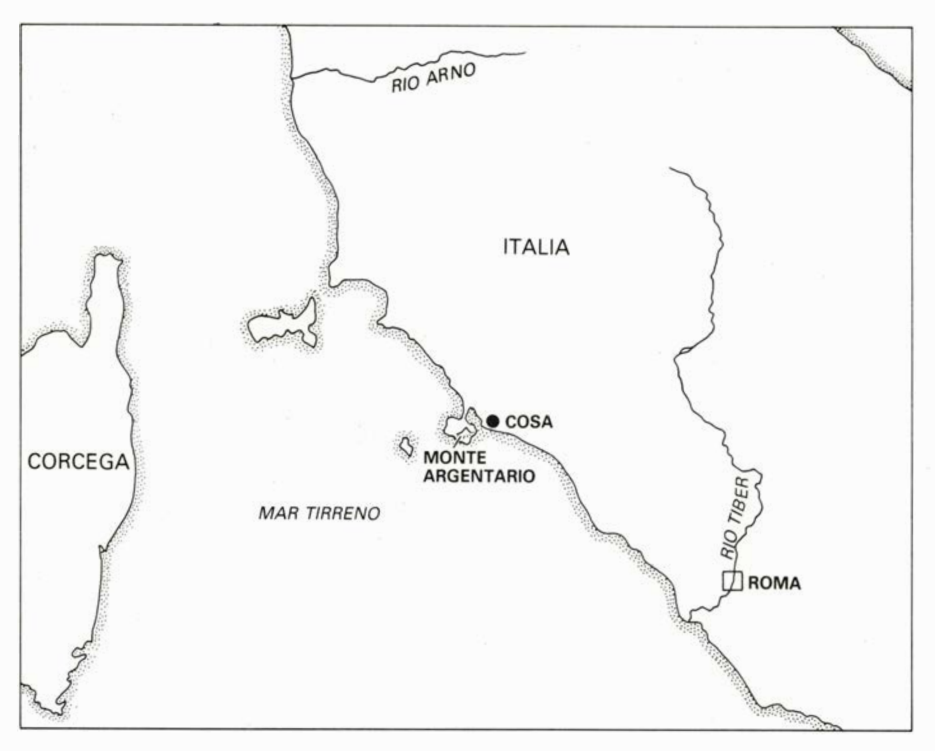
El alfarero fabrica primero un patrón, un molde sólido a escala natural de la lucerna, tallado en madera, hecho en piedra o moldeado en barro y cocido hasta alcanzar una gran dureza. Cualquier decoración que las lucernas fueran a ostentar, aparece sobre el patrón, como también las marcas que indican la posición del agujero de rellenar y el de la mecha. Utilizando el molde de patrón, el alfarero realiza entonces una matriz en dos partes; una parte reproduce en negativo la parte superior del patrón y la otra parte reproduce en negativo la inferior. La matriz de las dos partes se prepara de yeso o de barro cocido de gran dureza, y su relieve interior duplica a la inversa la forma y la decoración del original. La clave para la producción en masa reside en que el alfarero puede hacer tantas matrices cuantas quiera.

La etapa siguiente del proceso, que no requiere habilidades especiales, consiste en revestir las dos mitades de cada matriz con una pequeña capa de barro blando y retirar el resto. Las dos mitades se ponen aparte hasta que el barro que se está secando dentro de ellas adquiera la dureza del cuero. El barro moldeado se retira entonces de sus respectivas mitades y se junta con una delgada masilla de barro que sirve de cemento. El agujero de rellenar y el de la mecha de la lucerna entera son atravesados en los puntos indicados y la lucerna se deja aparte para endurecerse antes de ser cocida.

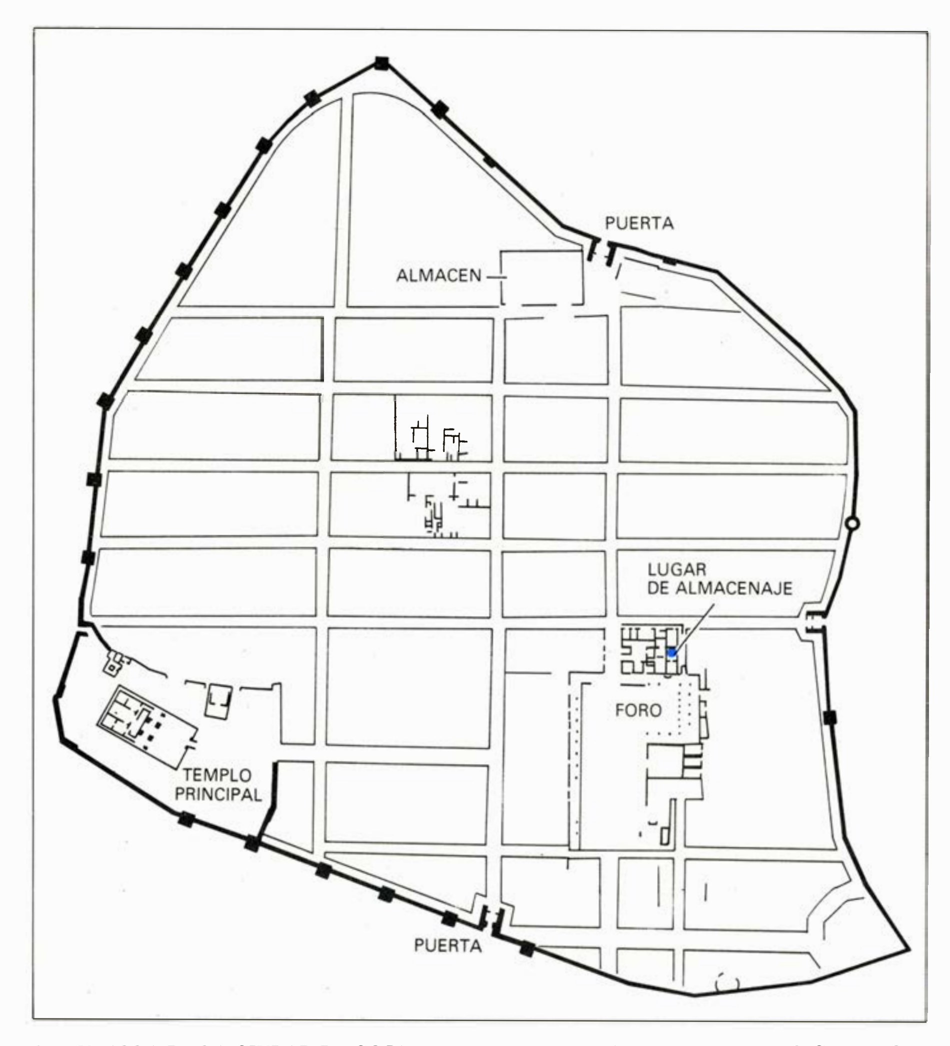
Todos los ejemplares, enteros o fragmentarios, hallados en Cosa son de terracota, menos uno. La excepción es una lucerna hecha de hierro. Todas menos una salieron de la rueda de alfarero o fueron moldeadas; la excepción es



1. EVOLUCION HIPOTETICA de las lucernas del Mediterráneo. Comienza (*1*) con una mecha que flota en un plato lleno de aceite. Algún improvisador apretó entonces una parte del borde del plato antes de cocerlo (*2*), proveyendo un soporte para la mecha. Las innovaciones posteriores (*3*) incluían una extensión de ese soporte, dando al plato una forma más ovalada, la adición de un asa al otro lado del soporte y un pliegue del borde del plato para evitar el derrame del líquido. El avance siguiente fue similar a la adición del asa; sobre el borde apretado se tendió un puente para crear una boquilla (*4*), en el lado opuesto del asa. Después, una boquilla formada con antelación tomó el lugar del puente agujereado (*5*). Esta boquilla pegada se hizo más voluminosa, el depósito más profundo, la abertura superior más pequeña y el asa vertical. En el siglo III a.C. alfareros de Asia Menor añadían lucernas hechas en moldes a su repertorio de las sacadas de la rueda de alfarero. Esta práctica económica no tardó mucho en extenderse por toda la cuenca del mar Mediterráneo.



2. LA CIUDAD DE COSA, un fuerte fronterizo, se estableció cerca del monte Argentario, en el año 273 a.C., después de que Roma y sus aliados conquistaran dos ciudades-estado etruscas de las cercanías.



3. MURALLA DE LA CIUDAD DE COSA, cuyo perímetro era más o menos pentagonal. Cercaba dentro de su circunferencia de 1,5 kilómetros un espacio total de 13¼ hectáreas. El abastecimiento de agua de la ciudad dependía de la lluvia y de fuentes ubicadas a extramuros. (El sustrato de Cosa era de piedra caliza.) El punto en color señala el almacén derrumbado donde se encontró la mayoría de las lámparas.

una lucerna hecha a mano. Las lucernas hechas sobre ruedas de alfarero, como cabría esperar desde el punto de vista de la historia de la evolución de estos útiles, caracterizaron los primeros siglos de la ocupación de la ciudad. De hecho algunas de ellas probablemente fueron traídas al lugar por los primeros pobladores. Esta conjetura queda apoyada por la localización del sitio donde se recuperaron: un hoyo sagrado bajo el suelo de la sala central del templo principal, el espacio consagrado a Júpiter. Era habitual que en la ceremonia romana, con ocasión de la fundación de una nueva ciudad, se cavara un hoyo de esas características en el lugar del futuro templo y se invocara la bendición de Júpiter. Entonces, los pobladores se alineaban y tiraban diversas pertenencias al hoyo como ofrendas al dios. El hoyo se llenaba, se erigía un altar provisional sobre él y el trabajo más lento de construir el propio templo comenzaba después. Por tanto, las lámparas encontradas en el hoyo no pueden datar de más tarde que la fundación de Cosa, en 273 a.C.

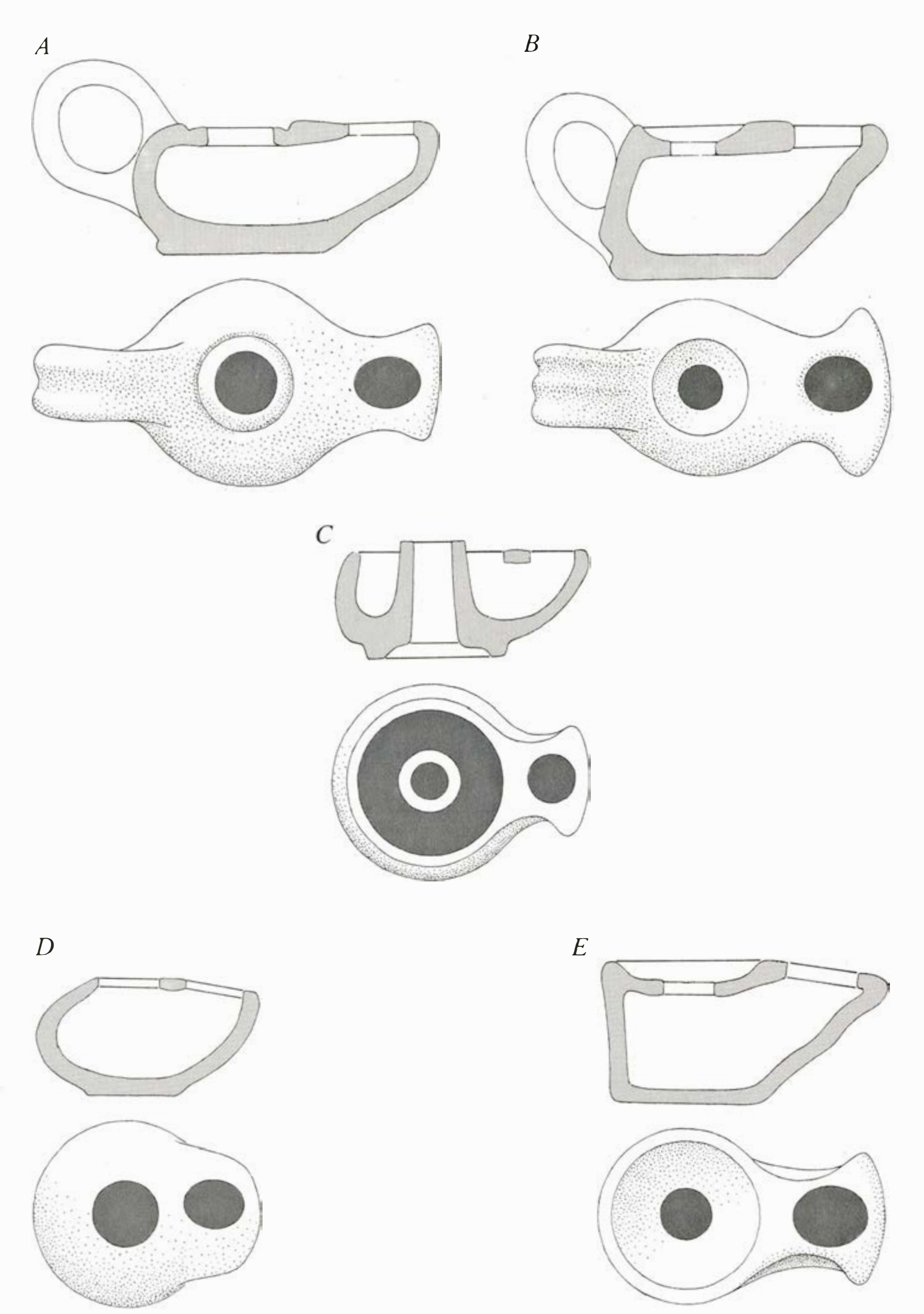
Los tres tipos de lucernas del hoyo constituyen una mayoría sobre los cinco tipos trabajados sobre ruedas de alfarero desenterrados en Cosa. Ahí están el tipo A, conocido como lucerna de "buñuelo", el tipo B, una lucerna de "cono truncado", y el tipo C, una lucerna "tubular". El cuarto tipo de lucerna hecho sobre rueda de alfarero y encontrado en Cosa –el tipo D, una lucerna "reloj"– aparece por primera vez en contextos estratigráficos más tardíos (fechados entre 225 y 200 antes de Cristo). El quinto tipo –el tipo E, una lucerna "cilíndrica"– aparece por primera vez en algún momento entre 120 y 70 a. C. La lucerna del tipo C era un compuesto: el cuerpo se hacía a imitación de un estilo griego más temprano, pero la boquilla era romana. Su forma era la adecuada para descansar sobre un soporte idóneo o para apoyarse sobre cualquier superficie llana. Pero su parte superior abierta constituía una invitación para ratas hambrientas de aceite, e insectos atraídos de noche por su llama debieron haber caído al depósito de aceite.

Después del año 70 a.C., ya no se encuentran en Cosa lucernas del tipo C y del tipo D; presumiblemente los alfareros de los entornos de Roma dejaron de hacerlas. Los otros tres tipos salidos de la rueda de alfarero, en cambio, continuaron fabricándose incluso después de que las lucernas moldeadas comenzaron a dominar el mercado. De hecho, se encontró una lucerna del tipo

A en Cosa en un contexto que la sitúa entre el año 250 p.C. y el tiempo del abandono de la ciudad en el año 500 de nuestra era.

La historia de Cosa, que transcurrió a la largo de casi ocho siglos, no fue parca en acontecimientos. Incluye varios intervalos en los que ni las lucernas de alfarerías regionales de Italia ni las importadas de ultramar llegaban a la ciudad. Las tres guerras púnicas fueron algunos de esos intervalos. La primera de las guerras (264-241 a.C.) comenzó una decena escasa de años después de la fundación de Cosa, y durante la segunda (218-201 a.C.) las tropas de Anibal estuvieron paseándose por Italia central. Lo que pudo haber sido un siglo de paz entre la derrota final de los cartagineses (149-146 a.C.) y las guerras que estallaron después del colapso del triunvirato romano (49-31 a.C.) fue interrumpido para Cosa por un ataque de piratas que capturaron y saquearon la ciudad, aproximadamente en el año 70 a.C.

Un tal Quintus Fulvius, un ciudadano próspero, parece haber estado avisado con anterioridad. En cualquier caso, enterró bajo el suelo de su despensa 2000 denarii, de los cuales el más tardío se acuñó en 72-71 a.C. Diríase que Fulvius murió en el saqueo, porque su tesoro permaneció imperturbado hasta que los excavadores de la Academia Americana lo descubrieron hace algunas campañas. Con el triunfo de Augusto en el año 31 a.C. volvió la paz a Roma, y unos años más tarde Cosa fue repoblada. La ciudad iba a perdurar otros quinientos años, aunque decreciendo gradualmente en habitantes e importancia hasta que terminó por ser abandonada.

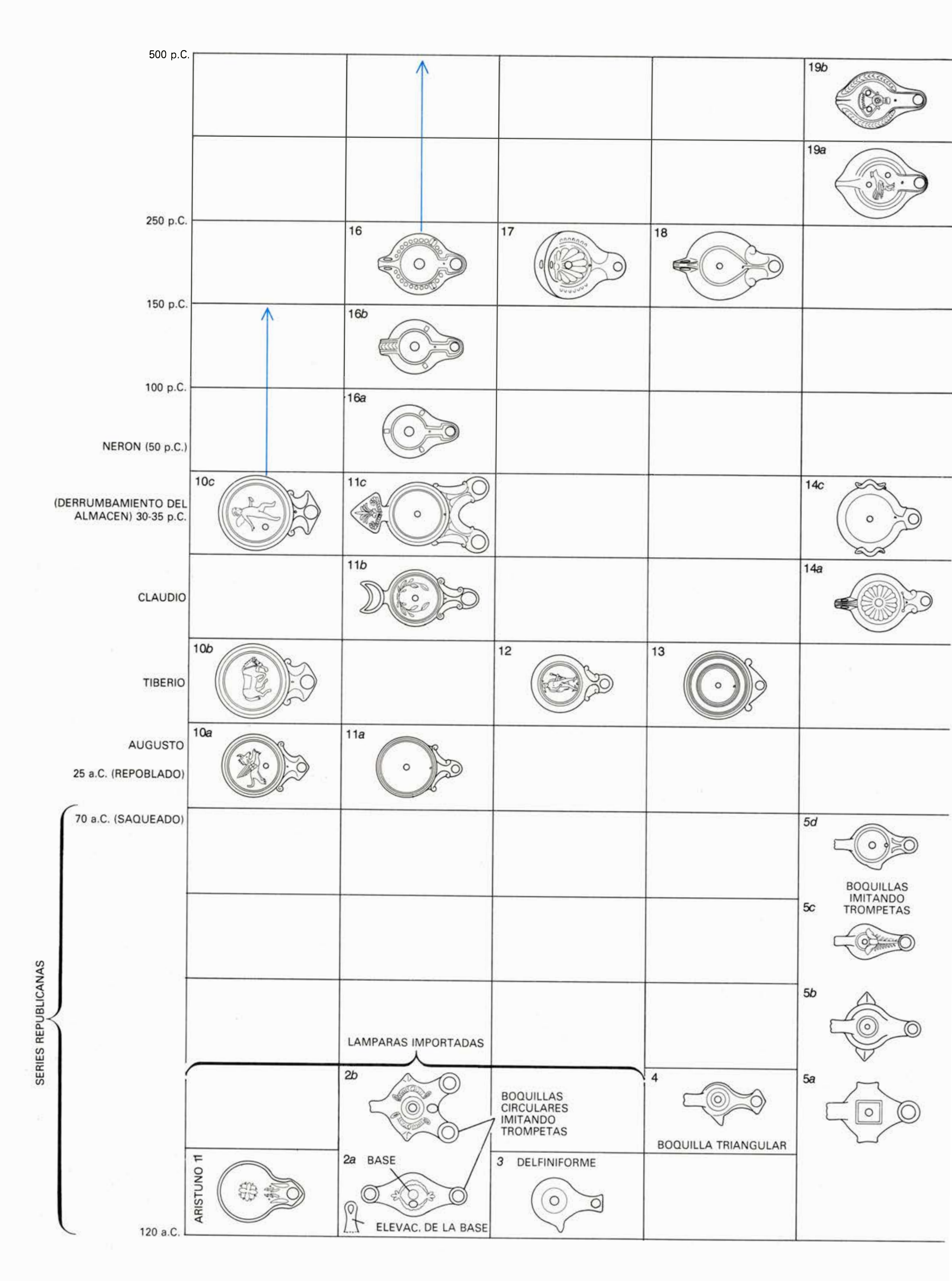


4. TODAS LAS LUCERNAS DE COSA PRIMITIVAS salieron de la rueda de alfarero. Tres de los tipos (*A, B, C*) se fabricaron, verosímilmente, antes del 273 a.C., año en que se fundó la colonia; estaban entre las ofrendas votivas hechas a Júpiter por los pobladores durante las ceremonias de fundación. Los dos primeros tipos tenían asas. El tercer tipo, de estilo compuesto, carecía de ella; su cuerpo imitaba un modelo griego más antiguo, pero su boquilla era de estilo romano. El cuarto tipo (*D*) apareció en Cosa entre los años 225 y 200 a.C. Igualmente sin asa, éste y el tipo *C* desaparecieron después del año 70 a.C. El quinto tipo (*E*) se presentó entre los años 120 y 70 a. C., cuando por primera vez se recuperan lucernas hechas con molde. Los tipos A, B y E continuaron usándose tiempo después de introducirse las lucernas moldeadas. Un ejemplar del tipo A surge de un contexto fechado entre los años 250 y 500 p.C.

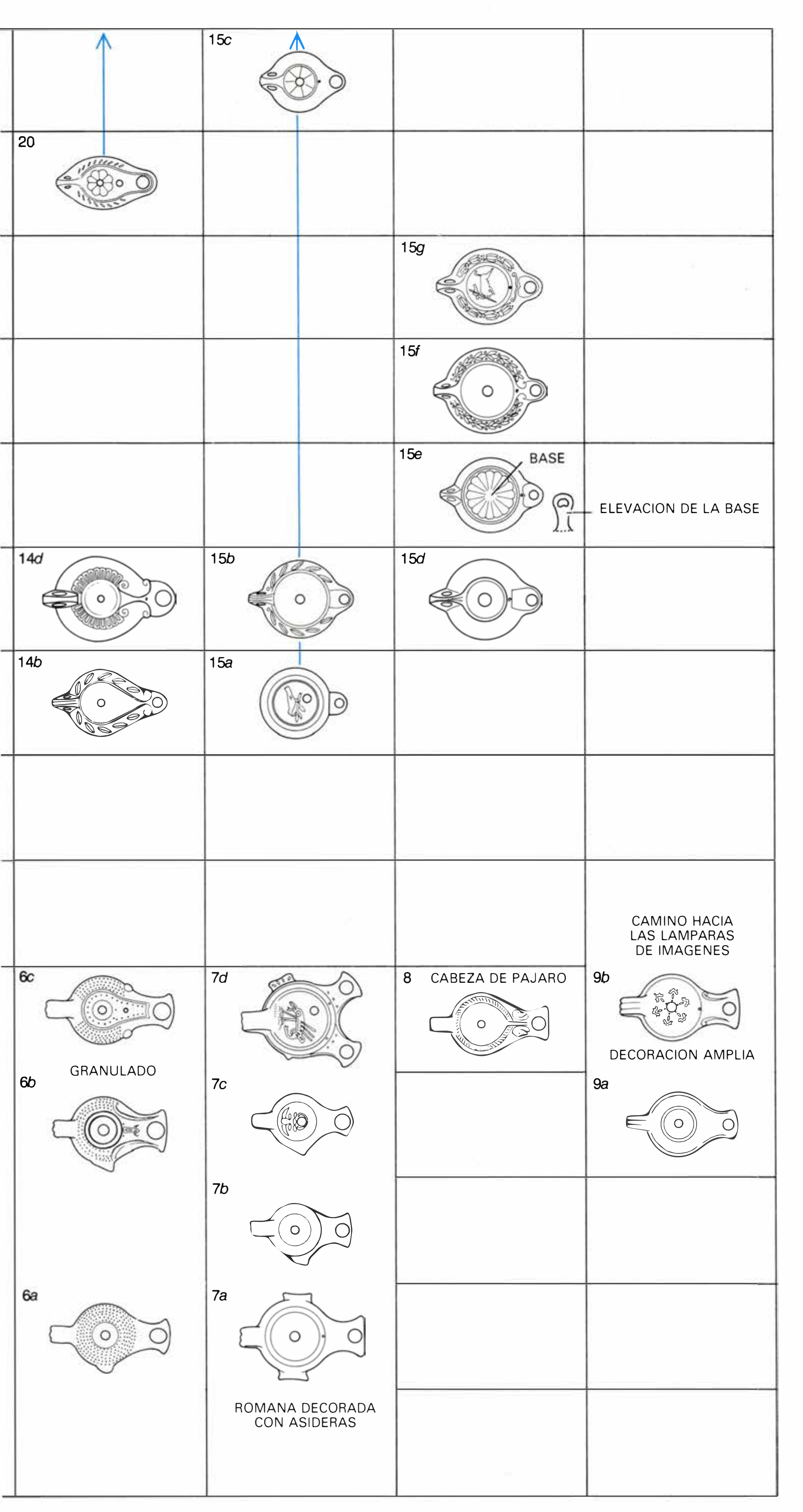
Las primeras lucernas hechas a molde no alcanzaron Cosa hasta después del año 120 a.C., pero entre esta fecha y el saqueo de los piratas la población de la ciudad no sólo se hizo con lucernas moldeadas de fabricación romana, sino también con otras moldeadas importadas de Sicilia, Grecia y Asia Menor. Con la técnica de la producción con moldes se incorporó la práctica de decorar las lámparas con bajorrelieves. Los temas para las decoraciones se inspiraron en representaciones artísticas de la época, incluidas muchas figuras de la mitología. A diferencia de las decoraciones de las lucernas fabricadas fuera, las decoraciones aplicadas por los alfareros romanos comenzaban a reflejar la influencia artística de los objetos que los ejércitos victoriosos habían traído como botín de las culturas superiores de Egipto, Grecia y Asia Menor.

Entre las importadas se han recuperado varias alrededor del templo principal de Cosa, fabricadas probablemente en Sicilia, y están datadas entre 120 y 70 a.C. Son conocidas como lucernas delfiniformes, porque su forma recuerda a un delfín. De pasta dura y gris oscura, tienen un barniz negro. Las decoraciones en relieve en su mitad superior representan uvas enlazadas con flores o filas concéntricas de una sola hoja y una flor de loto alternando con un diseño estrecho de abalorios y cilindros. Aunque importadas, tuvieron que fabricarse para el mercado romano, pues sus boquillas son de estilo romano.

Una lucerna griega encontrada en Cosa lleva el nombre de un alfarero ateniense bien conocido: Aristuno. El y su taller desaparecieron después del saqueo de Atenas por el general romano



500 p.C.
250 p.C.
150 p.C.
100 p.C.
NERON (50 p.C.)
(DERRUMBAMIENTO DEL ALMACEN) 30-35 p.C.
CLAUDIO
TIBERIO
AUGUSTO
25 a.C. (REPOBLADO)
70 a.C. (SAQUEADO)
SERIES REPUBLICANAS
120 a.C.
19b
19a
16
17
18
16b
16a
10c
11c
14c
11b
14a
10b
12
13
10a
11a
5d
BOQUILLAS IMITANDO TROMPETAS
5c
5b
LAMPARAS IMPORTADAS
2b
BOQUILLAS CIRCULARES IMITANDO TROMPETAS
4
BOQUILLA TRIANGULAR
5a
ARISTUNO 11
2a BASE
ELEVAC. DE LA BASE
3 DELFINIFORME



Sila en el año 86 a.C. Quizá la lucerna formó parte del botín de Atenas, traída por uno de los soldados de Sila, o tal vez la trajo a Cosa un viajero que había ido a Atenas antes que Sila. En cualquier caso, su mitad superior está decorada con un delicado diseño de escalas superpuestas.

De las lucernas hechas con molde en Cosa, una de las primeras y más finas es la de "puntos en relieve", llamada así porque la decoración consiste en filas concéntricas de granulaciones diminutas que cubren todo el cuerpo de la lámpara. Evidentemente, el patrón original de ese diseño moldeado era una lucerna metálica, quizás hecha de bronce. Se han encontrado ejemplares que son casi gemelos a estos productos romanos en Cartago, en el norte de Africa, y en la isla griega de Delos. Este tipo fue popular en Italia central hasta la época de Augusto. El patrón original debió desaparecer, porque se fue deteriorando la calidad. En las lámparas más tardías, los puntos se hacen progresivamente mayores y se encuentran sólo en la parte superior.

En los años posteriores al saqueo de Cosa los alfareros de Italia central perfeccionaron gradualmente un diseño que se exportaría con gran éxito. Era la lucerna romana "de imágenes", caracterizada por la presencia de dibujos en bajorrelieve dentro de un círculo interior sobre la parte superior de la lámpara. Dos ejemplos de ese tipo encontrados en Cosa datan del repoblamiento de la ciudad en los años 25-20 a.C. o de un poco después. Ambas lucernas son de muy alta calidad, lo que sugiere que se había perfeccionado el diseño en una fecha incluso anterior. Una de ellas, quizá la más antigua de las dos, tiene una boquilla triangular con volutas alcanzando una cenefa compleja que enmarca la decoración pictórica. La otra lámpara presenta una boquilla más larga y delgada con una punta redondeada y volutas que se enrollan hacia ambos extremos. Ninguna de las dos lleva asas.

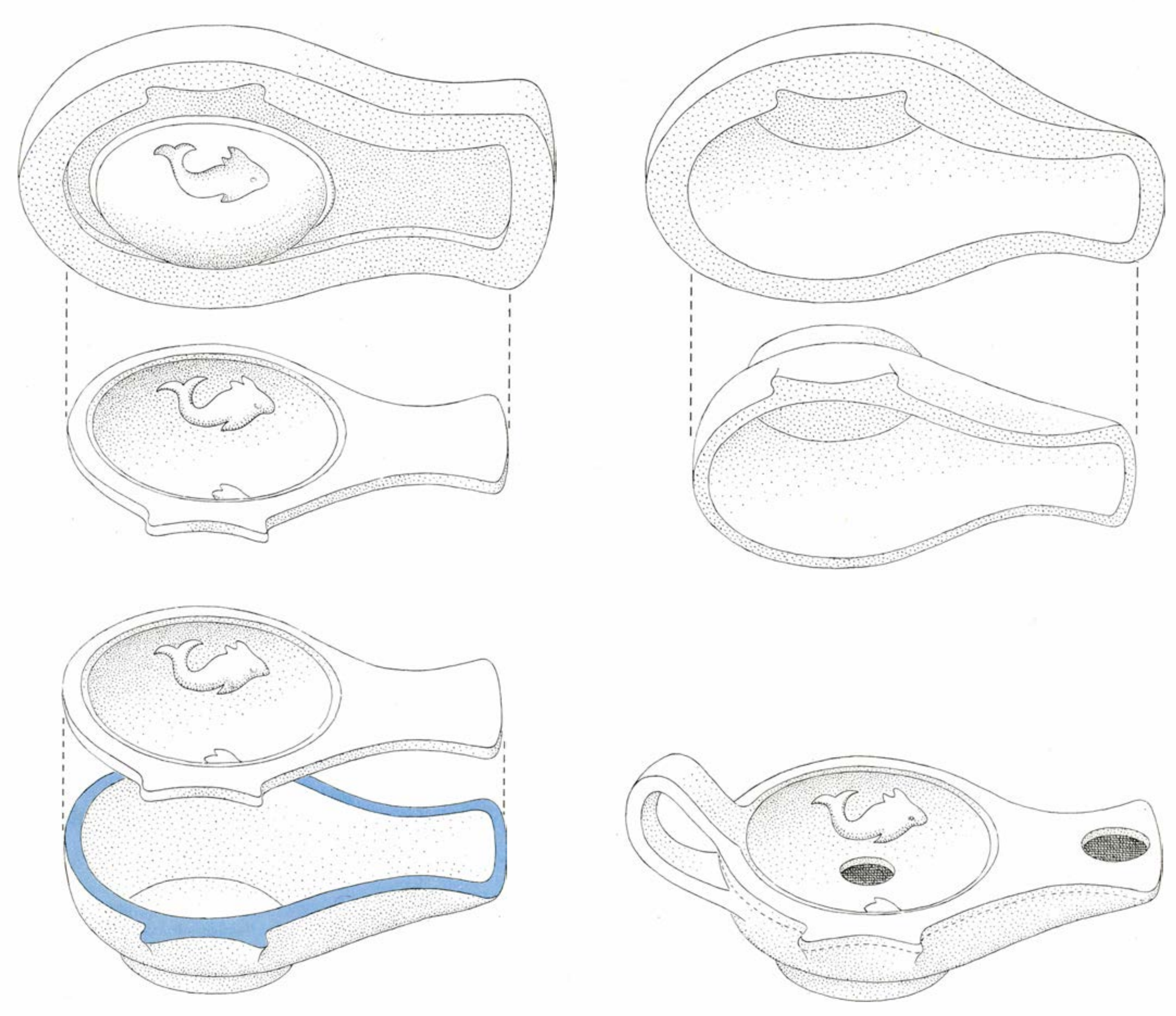
Las decoraciones en relieve del área

5. LUCERNAS HECHAS CON MOLDE descubiertas en Cosa. Pueden adscribirse a varios contextos históricos, comenzando con el período republicano (120-25 a.C.), abajo en el cuadro a la izquierda, y terminando en el año 500 p.C., arriba. Veinte variedades de lucernas están representadas por las 47 muestras dibujadas aquí. Tres, en las categorías 10, 15 y 16, continuaron fabricándose en el período indicado con las flechas. Hay cuatro variedades de importación (1, 2, 3 y 19).

central de las lucernas romanas "de imágenes", como sabemos por numerosos ejemplos encontrados en Italia y en otras partes, cubren una enorme gama de materias. Hay dioses y diosas, animales mitológicos como grifos e hipocampos y animales reales como gallos de pelea, caballos de carreras, caballos de circo, caballos con carro, caballos de tiro y caballos galopando sueltos. Aparecen escenas de Homero y Virgilio y de las hazañas de Hércules, de gladiadores en combate, temas cinegéticos e incluso escenas eróticas. Una representación popular, sobre lucernas empleadas como regalos de año nuevo, muestra una Victoria con alas sujetando un escudo inscrito con el mensaje *Annu faustu felice tibi* (un feliz y afortunado año para ti). Alrededor de la figura se disponen trozos de pan, pasteles, un tarro de miel, monedas y siempre una moneda mostrando al Jano de doble cabeza, el semidiós de los comienzos y de las puertas.

En algunas lámparas, la figura de Victoria no está rodeada de regalos, y su escudo lleva deseos para el bien común. Una de esas inscripciones dice *Fides publica* (la confianza pública), otra *Ob cives ser* (por la conservación de la ciudadanía). Incluso otras lucernas de imágenes llevan retratos de los emperadores romanos. El emperador Séptimo Severo, que gobernó desde 193 hasta 211 p.C., aparece con su esposa, Julia Domna, representados respectivamente como Sol y Luna (el sol y la luna).

La parte de Cosa que ha proporcionado mayor abundancia de lucernas es el conjunto de los lugares públicos, tiendas y puestos de almacenaje que constituyen un *atrium publicum*. (Ocho de esos atrios rodeaban el Foro.) Las primeras lucernas encontradas aquí datan aproximadamente del año 200 a.C. y las últimas aproximadamente del año 500 p.C. Pero el auténtico tesoro del atrium era el almacén de una ferretería. Lo destruyó poco antes del año 30 o 35 p.C. el derrumbamiento de un edificio adyacente. Cuando los arqueólogos de la Academia despejaron el área en la campaña de 1950, descubrieron un asombroso número de lucernas rotas. No sólo ha sido posible recoger ejemplares enteros entre los escombros, si-



6. PRODUCCION EN MASA facilitada por la adopción de moldes para formar las mitades superior e inferior separadas. A la izquierda se muestra una robusta matriz con su decoración central en negativo; debajo, la parte superior formada por revestimiento de una capa delgada de barro en el interior de la matriz. A la derecha, una matriz similar para la parte inferior; debajo, la parte inferior de la lámpara, lista para ser juntada con la parte superior. Abajo (*a la izquierda*) ambas mitades a punto de unirse; una masilla de barro (*color*) es el cemento. A la derecha, el producto acabado, con su boquilla.



7. DECORACIONES DE "LUCERNAS DE IMAGENES", típicas de los diversos temas representados en estos modelos populares moldeadas. Incluían, entre otras figuras mitológicas, a Cupido y grifos (*arriba*). Caballos y escenas circenses (*las dos inferiores*) también abundaban en las decoraciones populares.

no también, cuando la lucerna estaba decorada con un relieve, juntar los fragmentos suficientes para revelar gran parte o todo el decorado destrozado en ella.

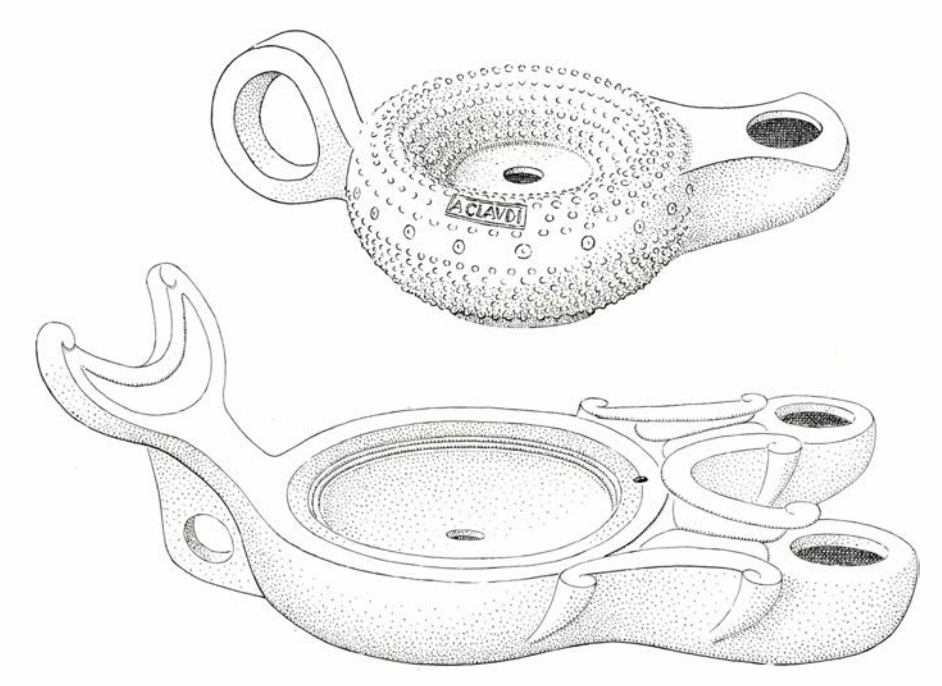
Los escombros del almacén proporcionaron numerosos ejemplos incluso de otros tipos de lucernas romanas. La primera corresponde al modelo de "boquilla triangular". Revela una evolución estilística desde las versiones tempranas con boquillas largas y estrechas hasta las lucernas con boquillas progresivamente más anchas. Un segundo diseño de boquilla estrecha, contemporáneo del primero, tenía al concluir su evolución una boquilla corta y fuerte con volutas de adorno que más bien parecen palomas enfrentadas.

Otra prueba de innovación en ese tiempo se puede encontrar en lucernas con boquillas largas y protectores del calor sobre sus asas. Muchas tenían boquillas dobles; había una con siete. Las lucernas parecen haber sido difíciles de fabricar y fáciles de romper; diríase que llevan al límite las posibilidades de una forma de lucerna que se pudo fabricar con una matriz de dos piezas. Y quizá quepa referirse a ellas como representantes del período barroco del diseño de lucernas romanas.

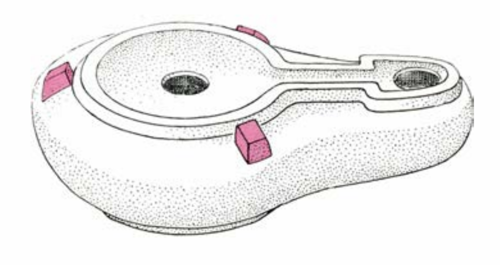
Los fragmentos del almacén revelan que un tipo de lucerna aún más conservador apareció en ese tiempo, tal vez en la última década antes del colapso del edificio adyacente. El asa de la lucerna formaba parte del diseño del molde y su hombro tenía una inclinación hacia abajo. Su boquilla presentaba una forma nueva: corta, redondeada y unida al hombro con volutas. Este tipo gozó siglos de popularidad y fue exportado (y a menudo copiado) a todo el Imperio. Al principio, su decoración en relieve estaba limitada, como lo era en las lámparas de imágenes más tradicionales, al área enmarcada de la parte superior, pero alrededor del año 150 p.C. la decoración se extendió también hacia el hombro inclinado.

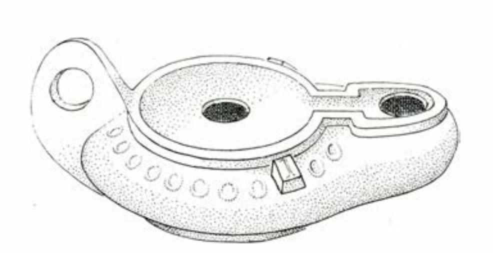
Un tipo de lucerna de molde completamente nuevo apareció en Cosa entre los años 50 y 70 p.C. Fue fabricado en el norte de Italia (uno de los talleres estaba en lo que hoy es Modena) y ha llegado a ser conocido como la "lucerna de fábrica". Según todos los indicios, su patrón debió tomarlo de una lucerna de bronce con tres enganches para la suspensión al mismo nivel, saliendo de su hombro, de modo que podía colgarse con cadenas desde el techo o desde un soporte en la pared. La matriz y las lucernas moldeadas reproducían al principio fielmente los enganches. Los productores de las lucernas de fábrica incluso agujereaban los enganches de barro durante el proceso de montaje, pero finalmente se prescindió de este toque de verosimilitud, y en cuanto pasaron los años, los enganches degeneraron en simples protuberancias redondeadas sobre el hombro.

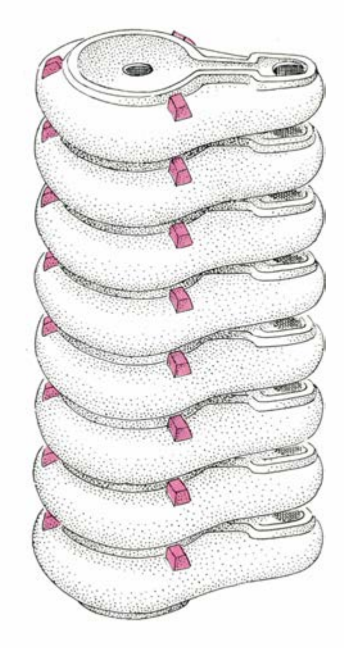
Esta sencilla y fuerte lucerna era fácil de embalar para su transporte naval (se encontró un cajón de estas lucernas todavía embaladas en Pompeya). Tam-



8. LUCERNA ANTIGUA hecha a molde; popular en Cosa, era un producto romano decorado con filas de puntos de relieve (*arriba*). Esta particular lucerna de "puntos de relieve" llevaba el nombre de un alfarero. La lámpara con las boquillas dobles (*abajo*) es uno de los nuevos tipos de lámparas romanas que aparecieron en los años 30-35 p.C. Una de las lámparas de boquilla múltiple tenía hasta siete.







9. "LUCERNA DE FABRICA", un producto barato, sencillo y fuerte hecho en el norte de Italia, que apareció en Cosa poco después del año 50 p.C. Su prototipo fue una lámpara de bronce con tres enganches suspensivos (*color*); éstas fueron en principio fielmente copiadas en las versiones en cerámica. Fácil de hacer y embalar para su transporte naval (*ver detalle del embalaje a la derecha*), las lámparas de fábrica dominaron el mercado por más de un siglo. La versión simple (*arriba a la izquierda*) adquirió más tarde un asa y decoraciones modestas (*abajo a la izquierda*). (Ilustraciones realizadas por Alan D. Iselin.)

bién era barata: el trabajo no especializado de esclavos podía llenar las matrices con barro y juntar las partes moldeadas una vez secas. Incluso el proceso de sacar las propias matrices pudo haber sido dejado a trabajadores no muy especializados. Como consecuencia, las lucernas de fábrica se produjeron por millares y conquistaron el mercado no sólo en el norte de Italia sino también en Italia central y en las provincias del norte y oeste del Imperio (por razones desconocidas no muchas alcanzaron las áreas del este y sur del Mediterráneo). Las lucernas de fábrica continuaron dominando el mercado por más de dos siglos. Antes de finales del siglo II p.C. se fabricaban también en Galia, Germania e incluso en Britania.

Mientras la mayoría de las lucernas de Cosa se realizaron en otras partes, algunas hechas con rueda de alfarero del tipo E parecen ser de fabricación local. Esta conclusión se funda en el tipo de barro empleado y en una marca de sello circular u oval en la parte superior de la boquilla. Dentro de la estampilla está o bien la letra *S* o bien las letras *SL*. Ocurrió que miembros de una familia llamada Sestius residieron en Cosa a lo largo de varias generaciones. Se sabe que fabricaron las tradicionales vasijas de almacenaje con dos asas, llamadas ánforas. Quizá cuando la demanda de ánforas era baja, el taller de la familia Sestius cambió hacia la producción de lucernas y de otra cerámica.

Solamente un nombre entero aparece en una de todas las lucernas encontradas en Cosa. En la parte superior de una moldeada con decoración de puntos en relieve (al contrario de en la parte inferior, como era costumbre) está impreso el nombre A CLAUDI. No hay ninguna clave para saber dónde se encontraba el taller de Claudi. No mucho antes del año 50 a.C., tres nombres comienzan a aparecer en las bases de las lucernas de Cosa, pero en formas abreviadas: COPPIRES (alternativamente COPRESTI), con la probable lectura de *Caius Oppius Restitutus*; LFABRICMAS (alternativamente LFABRICMASC), con la probable lectura de *L. Fabricius Masculus*; y un tercero, LMADIEC, con la probable lectura de *L. Munatius Adjectus*. La mayoría de las lucernas encontradas en Cosa, en cambio, no llevan marcas identificadoras.

Los últimos siglos de la existencia de Cosa fueron de lenta decadencia, decadencia desde los todavía bastante altos niveles mantenidos bajo el emperador Adriano (117-138 p.C.). A excepción de unas pocas y pobres copias de anteriores lucernas romanas, las de Cosa que datan de los años 250 a 500 p.C. son toscas importaciones del norte de Africa.

Treinta de las lucernas de este período tardío revisten particular interés histórico. Fueron encontradas en un recinto sagrado de Baco que sobrevivió décadas después de los edictos del emperador Teodosio por los que declaraba fuera de la ley al paganismo, en los años 391-92 p.C. El final del recinto sagrado y del culto a Baco en Cosa evidentemente fue violento. El mobiliario del recinto, que incluía una preciosa estatua del siglo I de Baco joven, vasijas rituales decoradas con serpientes reptando por las asas, y las 30 lucernas, fueron todos rotos y abandonados; alguien robó la cabeza de la estatua. Los escombros fueron cubiertos en principio por tierra traída por el viento y después por el techo y los muros caídos del recinto sagrado. Las monedas encontradas en las ruinas por los excavadores de la Academia sugieren que la destrucción tuvo lugar aproximadamente hacia el año 500 p.C.

¿Qué tipo de función se puede esperar de una lucerna? En diversas ocasiones festivas de Cosa, arqueólogos de la Academia han puesto a funcionar dos de las lámparas intactas salidas de la rueda de alfarero (una del tipo B y otra del tipo C). Hicieron mechas para las lucernas trenzando ocho hilos de algodón de una mezcla de una moderna linterna de pescador. Ambos tipos de lucernas ardían con una gran llama que acababa en un punto azul con poco o ningún humo. Las llamas eran más claras que las de una vela y duraban más o menos dos horas y media con un solo depósito de aceite de oliva.

Ahora bien, la colección de lucernas de Cosa no adquiere su valor por las funciones de las lámparas, ni por su fabricación o por sus cualidades artísticas, ni tampoco porque algunos ejemplares hallados en Cosa sean muy raros en otras partes del mundo romano. Su importancia le viene sobre todo del hecho de representar las modas e incluso las manías en la iluminación artificial de una pequeña y típica ciudad provinciana romana durante más de siete siglos. Se trata de uno de los intervalos temporales más dilatados conocidos hasta ahora por testimonios materiales en el área de la arqueología clásica. La colección de Cosa es una amplia representación de los productos de una industria internacional de la cuenca del Mediterráneo, que floreció, decayó y floreció de nuevo durante un período largo y significativo en la historia antigua. Finalmente, cuando el inventario de las lucernas de Cosa sea puesto a disposición de todos los estudiosos del mundo clásico, deberá ayudar a determinar las fechas de otras lámparas mediterráneas encontradas en contextos no tan claramente definidos.

Temas metamágicos

Algunos aspectos insólitos de la actividad matemática

Miguel de Guzmán

En muchas de las etapas más importantes de la filosofía, el pensamiento matemático ha sido considerado como modelo de conocimiento. La tendencia es bien patente en los pitagóricos, para quienes la misma constitución ontológica del universo es número y armonía. Así se expresa solemnemente Filolao en el siglo V a. C.: "Grande, todopoderosa, todoperfeccionadora y divina es la fuerza del número, comienzo y regidor de la vida divina y humana, participante en todo. Sin el número todo carece de fronteras y es confuso y oscuro. Porque la naturaleza del número proporciona conocimiento y es guía y maestra para todos en todo lo que es dudoso o desconocido. Porque nada de las cosas nos sería claro ni en su mismo ser ni en sus relaciones mutuas si no existiera el número y su esencia. Este es quien armoniza en el alma las cosas con su percepción, haciéndolas cognoscibles y congruentes unas con otras según su naturaleza, proporcionándoles corporeidad" (Diels, B.11).

Tras las fórmulas concisas que nos han llegado de los pitagóricos que, interpretadas al pie de la letra, pueden resultar un tanto estrafalarias, urge explicitar ese respeto y veneración por la forma y cuantificación, extraordinariamente modernos, que siguen siendo valederos para la ciencia actual. Esta actitud pitagórica ha perdurado a lo largo de la historia del pensamiento con notable firmeza. Platón es el intérprete, al mismo tiempo transmisor y purificador, del pensamiento pitagórico. Cuando el alma se queda a solas (*αύτη καθ'αύτήν*) ante el mundo inteligible (*κόσμοσ νοητός*) y trata de conocer el ser de las cosas, debe descomponer cada una de las ideas en sus elementos. Así llega a los elementos más simples (*άτομον ειδος*). Clasificada de esta manera, la idea compleja tiene una determinada clave numérica. Esta es la traducción platónica de la expresión más ontológica de los pitagóricos. El pitagorismo renace en el neopitagorismo del siglo I a. C. y en el neoplatonismo de los siglos III al VI p. C. Macrobio, entre los romanos, con su *Comentario al sueño de Escipión*, es uno de los transmisores de la Edad Media de esta corriente del pensamiento.

Galileo, a medio camino entre el Medievo y la Edad Moderna, se expresa así en su vejez: "Opino en verdad que el libro de la filosofía es el de la naturaleza, el cual se encuentra perpetuamente abierto ante nuestros ojos, pero como está escrito en caracteres diferentes de los de nuestro alfabeto no puede ser leído por todos. Los caracteres de este libro son triángulos, cuadrados, círculos, esferas, conos, pirámides y otras figuras matemáticas muy propias para una lectura de tal género" (*Carta a Licatti,* 1641).

Por su parte Descartes, a fin de avalar las afirmaciones del *Discurso sobre el método*, escoge una parcela del conocimiento privilegiada en que sus principios puedan aparecer más claramente ejemplificados. Esta es precisamente la geometría, junto con la dióptrica. La matemática, para él, constituye "un modo de habituar el espíritu a nutrirse con verdades y no contentarse con falsas razones".

La matematización del pensamiento como camino científico es hoy un dogma, a veces llevado a extremos ridículos, de la ciencia moderna. Kant llega a afirmar que "en cada una de las disciplinas de la naturaleza solamente se puede encontrar tanto de auténtica ciencia cuanto se encuentra en ella de matemática" (*Metaphysische Anfangsgründe der Naturwissenschaft. Vorrede*). Y la tendencia de todas las parcelas del conocimiento a fin de aumentar su credibilidad y su prestigio ha sido en los últimos tiempos, y sigue siendo, hacia el revestimiento, natural o forzado, de estructuras matemáticas.

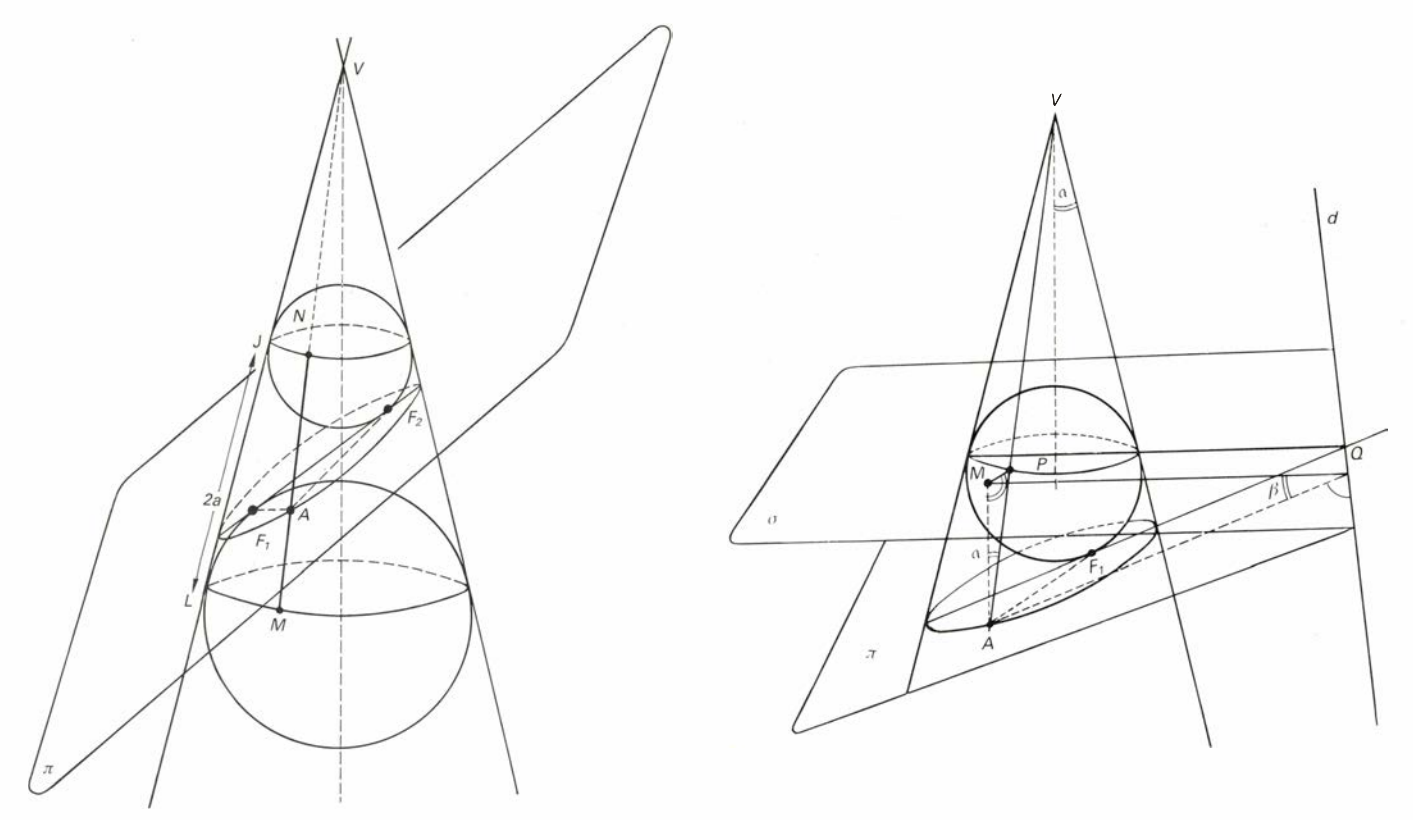
Puede uno preguntarse: ¿Merece la matemática este lugar privilegiado que se le ha atribuido de una forma tan constante? No han faltado filósofos que hayan considerado injustificada y aun nociva esta influencia invasiva del pensamiento matemático. Heidegger, con un juego de palabras, ha tratado de expresar la cuestión: la matemática "ist nicht *strenger*, sondern nur *enger*" (no es más exacta, sino sólo más estrecha). Y no sin cierta razón. El conocimiento humano contiene muchas más riquezas que las que el pensamiento matemático puede abarcar. Existen realidades profundas que el hombre, más o menos conscientemente, ansía aprehender cognoscitivamente que escapan a la matemática. Esta llega fundamentalmente a dominar la componente raciocinante del pensamiento sin tocar siquiera otras facetas del conocimiento intelectual. Desde el conocimiento matemático al conocimiento personal, que involucra a todo el hombre, existe ciertamente un abismo. El matemático y filósofo B. Pascal ha expresado así en sus *Pensamientos* la diferencia entre el espíritu matemático (*esprit de géométrie*) y el espíritu de discreción (*esprit de finesse*): "En el primero los principios son obvios pero alejados del uso común, de modo que cuesta girar la cabeza hacia este lado por falta de hábito. Pero por poco que se la vuelva se ven estos principios plenamente y sería necesario tener el espíritu totalmente falseado para razonar mal sobre principios tan toscos que es casi imposible que se escapen. Pero en el espíritu de discreción los principios son de uso común y están ante los ojos de todo el mundo. No hay que girar la cabeza ni hacerse violencia. No hay más que tener buena vista, pero es necesario tenerla buena, pues los principios son tan finos y numerosos que es casi imposible que no se nos escapen. Ahora bien, la omisión de un principio conduce a error. Por tanto es preciso tener la vista bien limpia para verlos todos y además un espíritu bien sano para no razonar falsamente sobre estos principios conocidos".

Y con todo, el pensar matemático merece un lugar privilegiado en el conocimiento por razón de su adecuación a su propio objeto, su evidencia y su certeza. No se trata de pensar con Leibniz que haya de venir el día en que dos filósofos de opiniones encontradas, en lugar de discutir sobre ellas, se sienten y digan: "Calculemos". Pero sí se puede afirmar que la matemática es un intento de creación de un universo para la satisfacción del ansia de conocimiento del hombre, hecho por éste a la medida de su propia mente. Intento no del todo logrado, pues, como veremos más adelante, este universo no carece de grietas

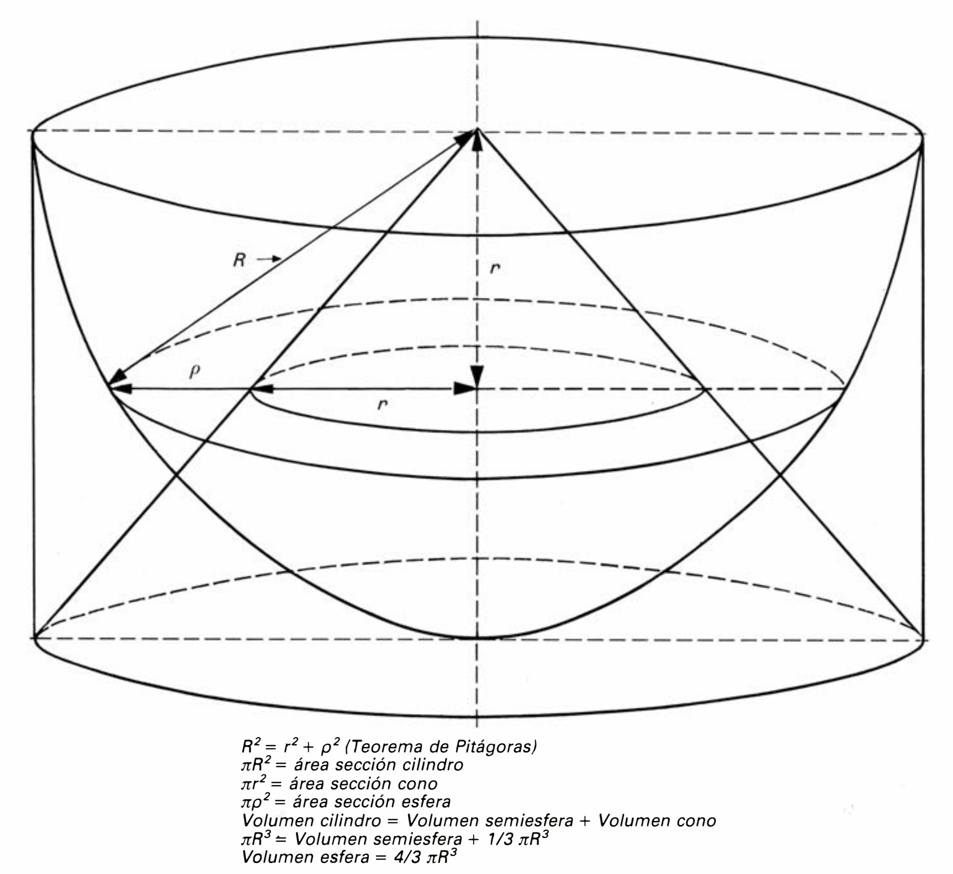
inquietantes. Esto lo hace de nuevo más cercano a la frágil condición de hombre, "esa caña pensante", como lo definió el mismo Pascal. Se diría que su seguridad le viene a la matemática de su carácter fundamentalmente tautológico, pero también es verdad que las tautologías que constituyen la rica estructura matemática representan un triunfo lleno de belleza y útiles implicaciones sobre el carácter reptante del raciocinio humano.

La perfección del pensamiento matemático ha llevado a considerarlo en muchas etapas de la historia de la humanidad como instrumento de comunión con la divinidad y con las fuerzas ocultas del mundo. Muchos de los monumentos megalíticos y ciclópeos, las pirámides de los egipcios, algunas de las construcciones de ciertas culturas precolombinas, la astrología de la antigüedad y de los tiempos modernos conservan entreverados elementos matemáticos. Los pitagóricos hicieron del cultivo a las matemáticas una secta religiosa con sus ritos y sus secretos. En la Biblia abundan los elementos numéricos con significados ocultos y la Cábala está plagada de ellos. En este aspecto ocupa un lugar importante el *Ars magna et ultima* de nuestro Ramón Llull (1235-1315), un sistema de conceptos y proposiciones primitivas que, por manipulación mecánica, darían lugar a todas las ciencias particulares, incluida la teología. Con ella se adelantó en varios siglos a la *Characteristica universalis* de Leibniz, que en su intento se confiesa inspirado por la obra de Llull. Entre los matemáticos más famosos no han faltado quienes han cultivado seriamente estos aspectos cuasimísticos de la matemática. En su *Harmonice Mundi* (1619), donde Kepler escribe su tercera ley sobre el movimiento de los planetas (el cubo de los ejes es proporcional al cuadrado de los períodos), se describen las relaciones armónicas matemáticas y musicales entre los planetas a través de las cuales Dios se da a conocer al oído espiritual de los hombres. Incluso Newton dedicó una buena parte de su energía a desvelar los misterios de la astrología de su tiempo, una faceta poco conocida de uno de los mayores matemáticos de todos los tiempos.

¿De dónde le viene a la matemática su carácter modélico? Su constitución puramente formal, sistema de postulados y leyes de inferencia con los que se van obteniendo los teoremas, la hace en cierto modo independiente de toda otra realidad que no sea la mente en su calidad raciocinante. De ahí su carácter incontrovertible y universal frente a opiniones, modas y progresos. Decimos "en cierto modo independiente", pues la elección de postulados e incluso de las leyes de inferencia es un acto que depende de la voluntad del matemático y de las convenciones de la comunidad matemática del momento. B. Russell ha dicho, aludiendo al carácter formal de la matemática, que se trata de "una ciencia en la que no sabemos de qué hablamos ni si lo que decimos es verdad o no". Pero esto es una verdad a medias. Lo que así se describe no es la matemática. Es solamente la estructura de su andamiaje lógico. La matemática nació como ciencia y sigue siéndolo y, por tanto, trata en sus desarrollos de desvelar y dominar alguna porción del mundo real, interior o exterior al hombre. Por eso no es un mero arbitrario juego



1. La elipse es la sección obtenida al cortar el cono circular recto por un plano π que forma un ángulo con el eje del cono mayor que el ángulo que forman las generatrices del cono con dicho eje. Trazando las dos esferas indicadas en la figura de la izquierda tangentes a π en F_1 y F_2 y a todas las generatrices del cono y trazando una generatriz r *cualquiera, se observa que $AF_1 = AM$, $AF_2 = AN$, $AF_1 + AF_2 = MN =$ constante. Así, para cada punto de la elipse la suma de sus distancias a los dos focos $AF_1 + AF_2$ es constante e igual al eje mayor de la elipse (***figura de la izquierda***). Asimismo, la figura de la derecha nos muestra, trazando el plano σ que contiene el círculo de la esfera superior, la recta* d *intersección de π y σ, la proyección M de A sobre σ y la proyección Q de M sobre* d, *que $AQ/AF_1 = AQ/AP = (AM/\cos\beta) / (AM/\cos\alpha) = \cos\alpha/\cos\beta =$ constante. Por tanto, para cada punto de la elipse la razón de distancias a una recta fija d y al foco F es constante (excentricidad)*



2. Modo en que probablemente Arquímedes de Siracusa (287-212 a. C.) descubrió el hecho de que el volumen de la esfera de radio R es $4/3\pi R^3$. Si se considera, como se indica en la figura, el cilindro recto circunscrito a la esfera y el cono con vértice en el centro de la esfera y base la del cilindro, y a continuación se observan las secciones de los tres cuerpos por un plano paralelo a la base común del cilindro y del cono, es fácil deducir, mediante el teorema de Pitágoras, que el área de la sección en el cilindro es la suma de las áreas de las secciones producidas en el cono y en la esfera. Como esto es válido para cada plano paralelo a la base se puede deducir, mediante el método de exhausción de Eudoxo (s. IV a. C.), que el volumen de la semiesfera es igual al del cilindro πR^3 menos el del cono $1/3\pi R^3$. Es decir, el volumen de la esfera es $4/3\ \pi R^3$

lógico. Sus postulados y sus leyes de inferencia están fuertemente inspirados en una fracción de la realidad. Es de aquí de donde le viene a la matemática su complejidad y su riqueza, reflejos de la riqueza y complejidad del mundo real mismo. El misterio de la adecuación de ese mundo matemático, tan propio de la mente que ha surgido de ella y de una mirada suya al universo, con la realidad externa es algo que no ha dejado de maravillar a los científicos de todos los tiempos. Este aspecto, origen de la aplicabilidad de las matemáticas, será considerado más adelante. También tendremos que considerar las serias limitaciones que en la estructura de la matemática han hecho patente los modernos progresos de la lógica reciente.

Pese a todo ello podemos afirmar que existen razones poderosas para considerar el conocimiento matemático como modelo de conocimiento científico, ya que ningún otro tipo de ciencia alcanza su objetivo propio con tanta eficacia, evidencia y certeza como lo logra el método matemático.

La afirmación de la naturaleza artística de la matemática puede sonar extraña en muchos oídos. Si arte es la producción por parte del hombre de un objeto bello, espero que esta afirmación resulte justificada al término de las notas que siguen. Desde que se empezó a analizar lo que es arte y belleza aparece explícita esta aseveración. Para los pitagóricos, la armonía, uno de los ingredientes de la belleza, va unida al número en la constitución ontológica de todo el universo. Aristóteles mismo se expresa así en su *Metafísica* (Libro XII, cap. III, v. 9): "Las formas que mejor expresan la belleza son el orden, la simetría, la precisión. Y las ciencias matemáticas son las que se ocupan de ellas especialmente".

Son muchos los testimonios que confirman la existencia de un verdadero placer estético en la creación y contemplación matemática. Así se expresa H. Poincaré en *La Valeur de la Science*: "Más allá de la belleza sensible, coloreada y sonora, debida al centelleo de las apariencias, única que el bárbaro conoce, la ciencia nos revela una belleza superior, una belleza inteligible únicamente accesible, diría Platón, 'a los ojos del alma', debida al orden armonioso de las partes, a la correspondencia de las relaciones entre ellas, a la euritmia de las proporciones, a las formas y a los números. El trabajo del científico que descubre las analogías entre dos organismos, las semejanzas entre dos grupos de fenómenos cualitativamente diferentes, el isomorfismo de dos teorías matemáticas es semejante al del artista".

Tal vez uno de los testimonios más elocuentes de esta afirmación sea el diario personal de Gauss. En este diario, escrito para él mismo especialmente en la etapa anterior a sus veinte años, período de muchos de sus grandes descubrimientos, va anotando, con un laconismo lleno de fuerza y de entusiasmo, sus observaciones sobre el universo matemático que se va desvelando ante sus ojos asombrados.

Pero este mismo placer estético en la contemplación matemática se da, en menor grado naturalmente, en todos aquellos a quienes se les presentan adecuadamente los hechos y métodos más salientes de la matemática elemental. Por supuesto que el goce estético de la matemática se encuentra en el mundo de la armonía intelectual, y así su percepción requiere una preparación inicial tanto mayor cuanto más elevado sea el objeto que se presenta. Por otra parte, así como el placer que puede proporcionar la pintura y la música, dirigidas a nuestros sentidos, al menos de modo inmediato, es perceptible hasta cierto grado con una contemplación más o menos pasiva, el placer estético de la matemática exige sin duda un grado de participación activa mucho más intenso. En el mundo de la matemática, a fin de gozar del objeto bello que se presenta, es necesario crearlo o recrearlo, de tal modo que el goce estético aquí presente es comparable más bien con el de hacer música, cantar, danzar, pintar, fabular...

Analicemos un poco más a fondo el origen de esta belleza matemática colocándonos desde una perspectiva clási-

ca. La belleza en general ha recibido muchas definiciones. S. Alberto Magno definió la belleza en el objeto como *splendor formae*, el resplandor del núcleo fundamental del ser, su unidad (armonía interna), su verdad (es decir su inteligibilidad y adecuación consigo mismo y con el mundo en su entorno), su bondad (su capacidad de llenar sus tendencias propias y las de los seres a su alrededor). Estas cualidades deben resplandecer de modo que sean accesibles y deleitables sin áspero trabajo. Otra definición clásica de la belleza que atiende a la componente subjetiva es la de S. Tomás de Aquino: *Pulchra sunt quae visa placent*. Bello es aquello que resplandece luminoso en su propio ser de modo que a quien lo mira le proporciona el sosiego y la facilidad de una percepción perfecta: la contemplación estética. Es aquello que se manifiesta de tal forma que produce una actividad armoniosa y compenetrada de las capacidades anímicas del hombre.

No se puede tampoco pretender describir la belleza matemática con un simple trazo. Me limitaré a señalar unos cuantos elementos de belleza que, a mi parecer, constituyen componentes bastante típicas en la actividad matemática. Un tipo de belleza matemática consiste en el orden intelectual que ante hechos aparentemente inconexos comienza a aparecer. Como un paisaje desde lo alto de la montaña que se desvela de una bruma que lo cubría. Todo el objeto contemplado aparece en conexión y la unidad lo invade. Entes aparentemente diversos que surgen en contextos diferentes resultan ser el mismo o estar ligados por una estructura armoniosa. La contemplación fácil de esta unidad inesperada es sin duda una de las fuentes de gozo estético de muchos hechos matemáticos.

Por poner un ejemplo tomado de la geometría elemental, consideremos cómo surgen las cónicas mediante la sección de un cono circular recto por un plano [*véase la figura 1*]. La sola contemplación de una figura bien trazada nos muestra enseguida que la elipse que así se define es el lugar de los puntos del plano cuya suma de distancias a dos puntos fijos es una longitud constante y que, al mismo tiempo, esta elipse es también lugar geométrico de todos los puntos tales que la razón de distancias a un punto y a una recta fija es igual a una constante, hecho nada sencillo de probar analíticamente.

Otro tipo de gozo matemático consiste en la realización de una ampliación de perspectivas con la que de una visión parcial se llega a la contemplación total de un objeto mucho más esplendoroso, en el que nuestro cuadro inicial queda englobado ocupando su lugar justo. En la exposición actual de la teoría de los números naturales, enteros, racionales, reales, complejos, se resume toda una aventura apasionante del espíritu humano que, a través de más de sesenta siglos de historia escrita, ha tenido sus callejones aparentemente sin salida, sus idas y venidas, sus paradojas.

Otro elemento estético presente muchas veces en la creación matemática consiste en la posibilidad de una contemplación descansada e inmediata de una verdad profunda, inesperada y llena de implicaciones. Como ejemplo de la matemática elemental puede citarse la demostración, que no requiere palabras, del teorema de Pitágoras que se presenta en la figura 5.

Para mí mismo, el haberme percatado hace unos años de la posibilidad de llevar a cabo la elegante demostración del teorema de Brianchon, de tanta profundidad e importancia para la geometría proyectiva, que se presenta en la figura 3, constituyó un motivo de goce estético enormemente intenso. No es que se pueda afirmar que todo el contenido de la matemática presente el mismo grado de belleza, ni siquiera que contenga belleza. Es indudable que el que una proposición sea cierta no implica que sea bella. Que $2^{32} + 1 = 641 \times 6700417$ es una verdad matemática interesante por motivos históricos, pero carente de gran belleza intrínseca. Existen teoremas claramente feos. Muchas teorías, en su nacimiento penoso y reptante, han resultado en un principio confusas y desprovistas de unidad y belleza. El cálculo infinitesimal de los tiempos de Newton y Leibniz constituye probablemente uno de los logros más importantes y útiles de la ciencia moderna, pero el grado de confusión y fealdad en que en un principio se encontraba contribuyó bastante a que su expansión y aceptación fuesen mucho más lentas de lo que la teoría merecía.

¿Qué características debe ofrecer un hecho matemático para que se pueda calificar como bello? La belleza matemática parece incluir cualidades tales como seriedad, generalidad, profundidad, inevitabilidad, economía de pensamiento, transparencia, sobriedad, adecuación... La seriedad se manifiesta en las ideas que pone en conexión, que normalmente dan lugar, en su desarrollo, a una buena porción del campo matemático en que tal hecho se encuentra, ya sea porque el método que lo crea es la clave que impregna dicho campo, ya sea porque el hecho en cuestión es el germen de todo ese cuerpo matemático. La generalidad se ha de dar con una cierta mesura. La generalización por sí misma no es en muchos casos más que el producto de una manía, sin gran valor; pero es cierto que un hecho demasiado concreto no despierta una gran admiración. Los matemáticos suelen calificar un método de "elegante" para indicar el tipo de sobriedad, economía de medios y transparencia que a veces se encuentra en la demostración de tal o cual teorema o hecho matemático. El proceso diagonal de Cantor, el método de dualidad en geometría proyectiva son ejemplos característicos de esta cualidad. Allí donde hay belleza matemática, ésta no se agota y su contemplación nunca deja de producir ese sentimiento de satisfacción, adecuación y acabamiento que una obra arquitectónica perfecta produce en el ánimo de quien la contempla.

La cualidad artística de la matemática se manifiesta asimismo en el proceso de su creación, que participa mucho de las características del proceso creativo de cualquier otro arte. Existe un magnífico estudio psicológico de J. Hadamard, gran matemático él mismo, sobre el proceso creativo en matemáticas (*The psichology of invention in the mathematical field*, Princeton University Press, 1945). Otro de los clásicos sobre este tema es un artículo famoso, muchas veces reproducido, de Poincaré (*L'invention mathématique*). Quien no haya tenido alguna experiencia creativa en matemáticas no podrá menos de sentirse asombrado ante las observaciones de Poincaré sobre el proceso matemático. A juzgar por el papel que desempeña la intuición, la inspiración, el trabajo y el descanso, y aún el sueño, o el ensueño, uno pensaría asistir a la composición de una sinfonía musical. Y en este sentimiento de dádiva repentina que la creación matemática comporta a menudo coinciden muchos matemáticos famosos como atestigua el mismo estudio de Hadamard. Por eso puede afirmar Poincaré con toda razón: "Puede extrañar el ver apelar a la sensibilidad a propósito de demostraciones matemáticas que, parece, no puede interesar más que a la inteligencia. Esto sería olvidar el sentimiento de belleza matemática, de la armonía de los números y las formas, de la elegancia geométrica. Todos los verdaderos matemáticos conocen este sentimiento estético real. Y

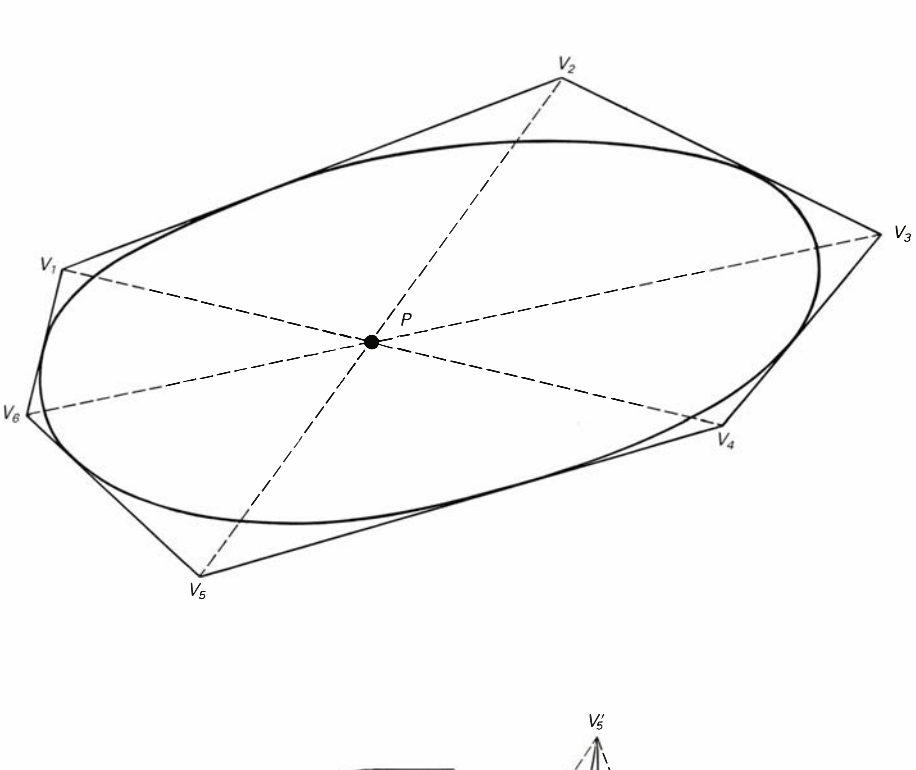
ciertamente esto pertenece a la sensibilidad. Ahora bien, ¿cuáles son los entes matemáticos a los que atribuimos estas características de belleza y elegancia y que son susceptibles de desarrollar en nosotros un sentimiento de emoción estética? Son aquellos cuyos elementos están dispuestos armoniosamente, de forma que la mente pueda sin esfuerzo abrazar todo el conjunto penetrando en sus detalles. Esta armonía es a la vez una satisfacción para nuestras necesidades estéticas y una ayuda para la mente, a la que sostiene y guía. Y al mismo tiempo, al colocar ante nuestros ojos un conjunto bien ordenado, nos hace presentir una ley matemática... Así, pues, es esta sensibilidad estética especial la que desempeña el papel de criba delicada de la que hablé antes. Esto permite comprender suficientemente por qué quien no la posee no será nunca un verdadero creador".

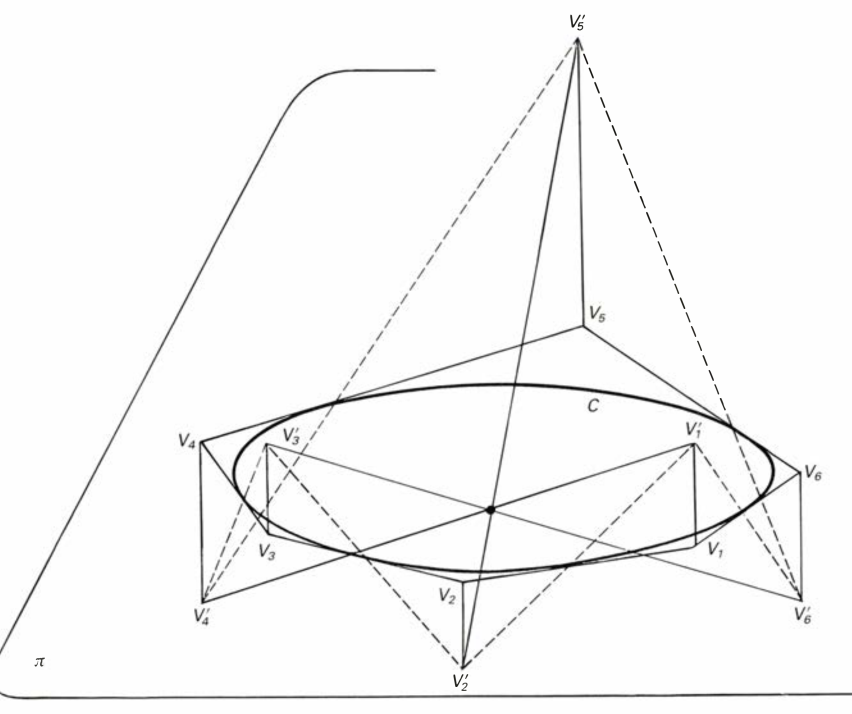
El que la matemática participe, efectivamente, de la condición de creación artística no da, por supuesto, carta blanca a los matemáticos para entregarse a un esteticismo estéril. La calidad artística de la matemática es como una dádiva con la que se encuentran quienes se dedican a esta actividad que es, al mismo tiempo y en grado muy intenso, ciencia y técnica. A este propósito resultan muy acertadas las sensatas observaciones de uno de los mayores matemáticos de nuestro tiempo, creador él mismo de un sinfín de campos matemáticos diversos. Así dice John von Neumann en su artículo *The Mathematician*: "A medida que una disciplina matemática se separa más y más de su fuente empírica o aún más si está inspirada en ideas que provienen de la realidad de un modo sólo indirecto, como de segunda o tercera mano, está más cercada por graves peligros. Se va haciendo más y más esteticismo puro, se convierte más y más en un puro arte por el arte. Esto no es necesariamente malo si el campo en cuestión está rodeado de otros campos relacionados con él que tengan todavía conexiones empíricas más cercanas, o si la disciplina en cuestión está bajo la influencia de hombres dotados de un gusto excepcionalmente bien desarrollado. Pero existe un grave peligro de que este campo venga a desarrollarse a lo largo de las líneas de menor resistencia, de que la corriente, tan lejos de su fuente, venga a disgregarse en una multitud de ramas insignificantes y de que la disciplina venga a convertirse en una masa desorganizada de detalles y complejidades. En otras palabras, a gran distancia de su fuente empírica, o bien después de mucha incubación abstracta, un campo matemático está en peligro de degeneración".

La matemática es también un instrumento de exploración de la naturaleza. Aristóteles ha expresado así el carácter "liberal" de la matemática en su mismo nacimiento: "Pero una vez que todas las técnicas necesarias se constituyeron, se vieron surgir ciencias cuyo objeto no puede ser ni la comodidad ni la necesidad. Nacieron primero en los climas donde el hombre puede entregarse más fácilmente al ocio y así las ciencias matemáticas nacieron en Egipto, donde la casta de los sacerdotes ocupaba de esta manera sus ocios" (*Metafísica*, Libro I, cap. I, v. 18).

La observación es interesante y expresa la concepción predominante en la Grecia clásica sobre el carácter desligado de toda consideración utilitaria de la matemática. Pero la afirmación de Aristóteles es falsa. La matemática occidental surgió primero entre los babilonios con fines prácticos bien concretos, económicos y astronómicos.

Las tablillas babilónicas que se conservan son un buen testimonio de ello. Hacia el siglo VIII a. C. se puede decir que la astronomía babilónica alcanza un grado de desarrollo capaz de medir



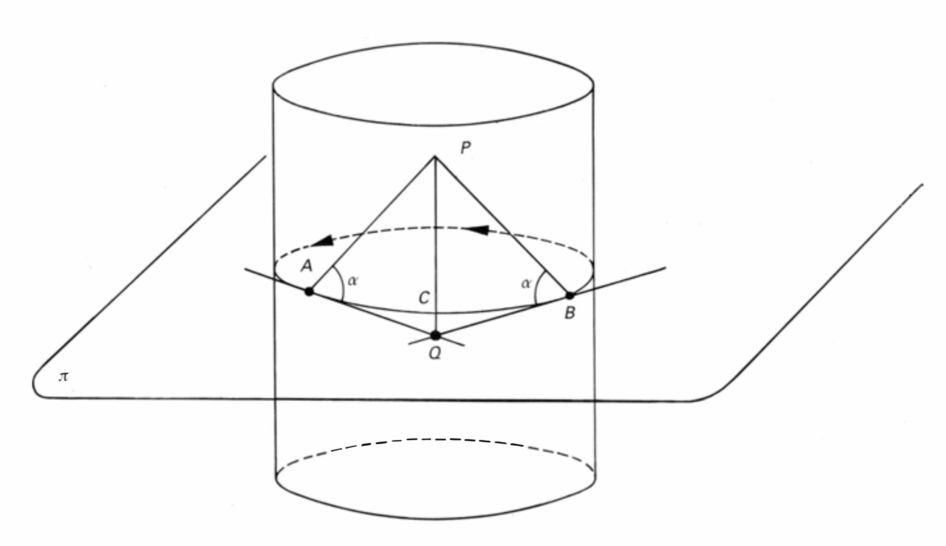


con bastante exactitud y predecir diversos fenómenos astronómicos. Se podría pensar que este tipo de aplicación práctica no es el mismo al que la técnica moderna nos tiene acostumbrados. La astronomía de aquel tiempo, podría suponerse, era especulativa. Pero no es cierto, los astrónomos babilonios seguían una finalidad práctica con sus mediciones y observaciones. Se trataba de elaborar con ellas los calendarios y vaticinios para sus monarcas. El carácter especulativo y contemplativo de las matemáticas predominó entre los griegos. La matemática pura tal como hoy la conocemos es creación de los matemáticos griegos de los siglos VI al III a. C. culminando su creación con los *Elementos* de Euclides, que constituyen una recopilación del saber profesional hasta entonces.

El carácter práctico y aplicado de la matemática aparece esporádicamente en la obra de Arquímedes, por ejemplo, quien llega a presentar sus invenciones casi como una aberración por la que pide disculpas. Así lo explica Plutarco en la vida de Marcelo, el general romano que conquistó la ciudad de Siracusa después de un largo asedio: "... aunque sus inventos le habían dado la fama de poseer una sagacidad más que humana no se dignó dejar tras de sí ningún trabajo escrito sobre tales materias, sino que, considerando como innoble y sucio el objeto de la mecánica y toda suerte de artes dedicadas a la utilidad y al provecho, puso toda su ambición en aquellas especulaciones cuya belleza y sutileza no están teñidas por ningún color que aluda a las necesidades ordinarias de la vida".

La matemática aplicada se refugia sobre todo en la astronomía, que en el período helenístico alcanzó un enorme grado de desarrollo, culminando en la obra de Ptolomeo, el *Almagesto*, compendio de toda la evolución milenaria de la astronomía. La razón profunda de esta situación se debe buscar en el hecho de que los fenómenos astronómicos resultan estar más cerca de la abstración, con su relativa simplicidad y pureza, comparados con los fenómenos inmediatos al hombre. Las órbitas de los planetas, casi circulares, su repetición rítmica, hacen de los fenómenos astronómicos algo muy cercano al experimento analítico, que ni babilonios, ni egipcios, ni griegos conocieron. Esta circunstancia motivó que, desde la antigüedad y a través de la Edad Media, hubiese una ruptura fundamental entre el mundo de los astros y el mundo sublunar. El primero, de naturaleza cuasidivina, sujeto a las leyes exactas y cuantitativas de la matemática. El segundo, enturbiado por la presencia de fuerzas confusas no fácilmente inteligibles ni comprensibles por el entendimiento humano. La ciencia moderna se caracteriza por la superación de esta repugnancia a considerar el mundo de los astros y el mundo sublunar como sujetos a las mismas leyes y a tratar de cuantificar los fenómenos terrestres, caída de los cuerpos (Galileo), atracción gravitatoria (Newton), etcétera. La conjunción de este cambio de mentalidad con la invención de un método tan poderoso como el análisis matemático (hechos de ningún modo independientes) dieron lugar al nacimiento, en primer lugar, de la física moderna y, más tarde, a la tendencia generalizada hacia la matematización de todas las disciplinas del saber humano.

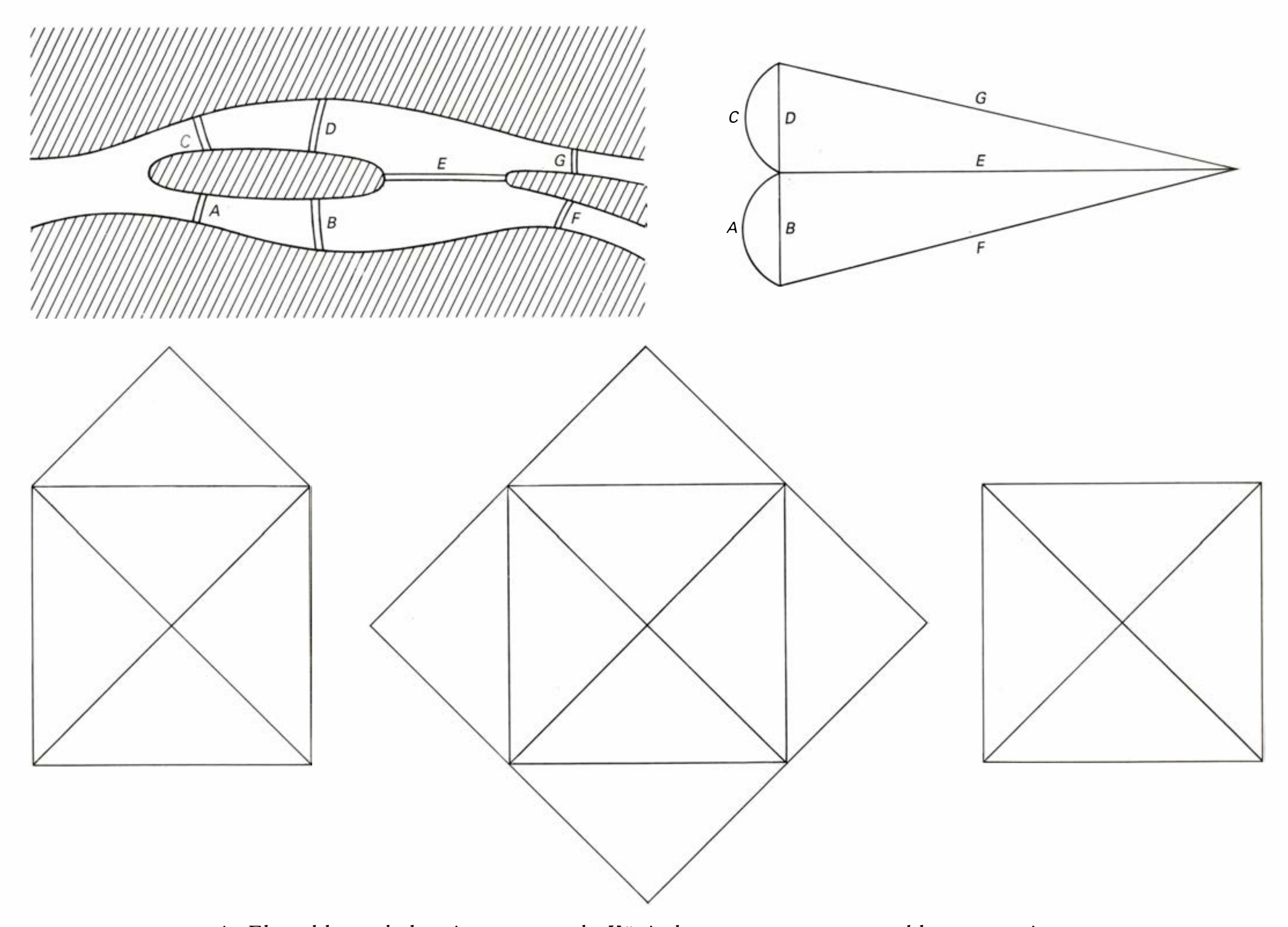
La aplicabilidad de las matemáticas se ha calificado como un misterio difícil de entender. Tal vez a un nivel elemental el hecho no parezca tan sorprendente. La aplicabilidad de la noción de número en la vida cotidiana del hombre, presente ya en las culturas más primitivas, parece un hecho que no necesita mucha explicación. Y sin embargo detrás de este hecho se esconde el mismo misterio que subyace a la extraña ade-



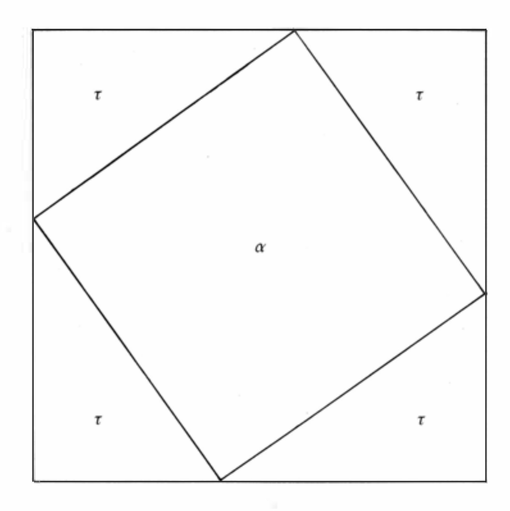
3. El hexagrama místico. El teorema de Brianchon afirma que si en un círculo (o en una cónica cualquiera) se circunscribe un hexágono arbitrario, las tres diagonales que unen vértices opuestos se cortan en un punto. La siguiente demostración (**dibujo central**)*, publicada por el autor por primera vez en 1976, es un buen ejemplo de las características del tipo de belleza que tantas veces se presenta en los hechos y métodos matemáticos. Se basa en las dos sencillas observaciones siguientes: (1) Si se fija un sentido sobre un círculo C en el plano π, se toman dos puntos A y B sobre C, se trazan semirrectas tangentes a C en A y B respectivamente en sentido contrario y finalmente se toman dos rectas hacia arriba tangentes al cilindro recto de sección C formando ángulos de igual magnitud con el plano π, entonces estas dos rectas se cortan en un punto P (o bien son paralelas si A y B son diametralmente opuestos). La proyección de P sobre π es el punto Q intersección de las dos semirrectas tangentes en A y B. Y recíprocamente, si se toman Q, se toma P situado en la vertical sobre Q y se une P con A y B, entonces los ángulos PAQ y PBQ son iguales. (2) Tres rectas en el espacio que se cortan dos a dos y no están en un mismo plano pasan las tres por un mismo punto del espacio. Con estas dos observaciones el teorema resulta de modo sencillo. Se considera el círculo C en el plano π y el hexágono circunscrito $V_1V_2V_3V_4V_5V_6$* (**figura de la derecha**)*. Por V_1 levantamos una perpendicular a π hasta un punto arbitrario V'_1. Unimos V'_1 al punto de tangencia de V_1V_2 a C y prolongamos hasta cortar la perpendicular por V_2 a π en V'_2. Procediendo así sucesivamente obtenemos en el espacio $V'_1V'_2V'_3V'_4V'_5V'_6$ y volvemos a V'_1. Consideremos ahora $V'_1V'_4$ y $V'_2V'_5$. Observemos que $V'_1V'_2$ y $V'_4V'_5$ son dos rectas en la situación de la observación (1). Así, o son paralelas o se cortan. Por tanto $V'_1V'_4$ y $V'_2V'_5$ están en un plano y, así, o bien son paralelas o se cortan. Se excluye que sean paralelas por construcción. Así se cortan. Lo mismo vale para $V'_2V'_5$ y $V'_3V'_6$. Se cortan. También $V'_1V'_4$ y $V'_3V'_6$ se cortan. Usando la observación (2) resulta que las tres rectas se cortan en un punto, pues no están en un mismo plano por construcción. Por tanto sus proyecciones sobre π también, pero éstas son precisamente las diagonales V_1V_4, V_2V_5, V_3V_6.*

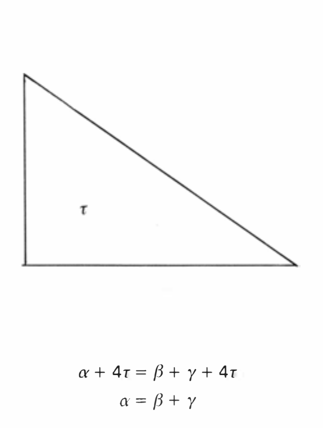
cuación a realidades más complicadas de las teorías matemáticas más complejas. El matemático observa el mundo real dentro de sí mismo y en su entorno, abstrae unas cuantas propiedades que a él le interesa destacar, con ellas construye su mundo mental abstracto que pone en movimiento mediante las leyes básicas de su mente. Guiado por su intuición y sentido estético dirige su construcción escogiendo unas cuantas de las infinitas posibles estructuras que podría crear. Descubre así un complejo mundo mental, que él ha perseguido por su coherencia, armonía, belleza intrínseca. Y de repente, o paulatinamente, según los casos, observa que cierta parcela del mundo real es iluminada y plenamente explicada mediante aquella teoría matemática que hasta entonces se había mostrado alejada de la realidad. Esta vivencia está presente en la historia de la matemática. El misticismo filosófico de los pitagóricos del siglo VI a. C. es probablemente una respuesta a ella, como lo es ciertamente el de Kepler en el siglo XVII. Apenas se puede uno imaginar la sensación de congruencia y belleza que éste debió de experimentar al comprobar que la teoría de las cónicas, tan cultivada y perfeccionada por sí misma desde los días de Apolonio de Pérgamo (s. III a. C.), se adaptaba plenamente a la explicación del movimiento planetario, viendo que elementos con significado en apariencia puramente matemático como los focos de tales cónicas adquirían una significación física tan importante (el Sol ocupa uno de los focos de cada una de las cónicas descritas por los planetas). El entusiasmo de Kepler ante esta coherencia, como en el caso de los pitagóricos, explica y justifica sus extrapolaciones místicas en su obra *Harmonice mundi*.

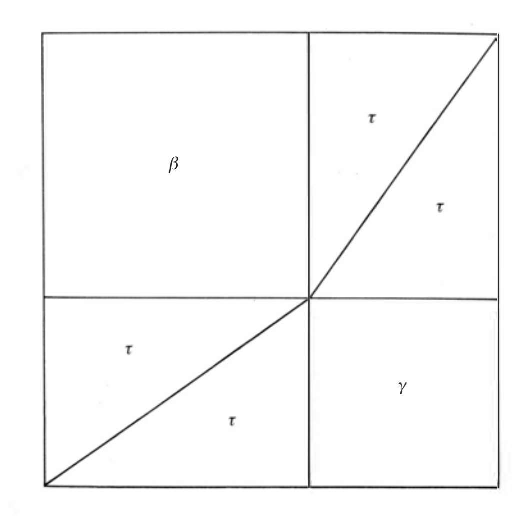
La razón más profunda de esta especie de armonía pre-establecida que se puede observar en la aplicabilidad de la matemática al mundo real está íntimamente ligada al problema filosófico fundamental de la teoría del conocimiento, en el que no qusiera introducirme aquí. Me limitaré a iluminar este punto con tres testimonios recientes. J. von Neumann escribe: “Yo creo que constituye una aproximación relativamente buena a la verdad –una verdad tan complicada que otra cosa que una aproximación a ella es impensable– el decir que las ideas matemáticas se originan en la experiencia, si bien su genealogía es a veces larga y oscura. Sin embargo una vez se han estructurado a partir de ella comienzan una vida peculiar, independiente, y se podría comparar el objeto matemático de la forma más adecuada con un sujeto creador, al que subyacen casi exclusivamente motivaciones estéticas, pero en ningún modo se puede comparar con una ciencia empírica”. En su artículo sobre *L'architecture des mathématiques*, N. Bourbaki afirma: “Que existe una relación íntima entre los fenómenos experimentales y las estructuras matemáticas parece confirmarse plenamente de la forma más inesperada mediante los descubrimientos más recientes de la física contemporánea. Pero no sabemos absolutamente nada sobre los fundamentos de este hecho (suponiendo que se pudiera encontrar realmente significa-



4. El problema de los siete puentes de Königsberg, su esquema y problemas semejantes







5. Teorema de Pitágoras

ción a estas palabras) y tal vez no lleguemos a saber nunca sobre ello". Por su parte, E. P. Wigner cierra su famoso artículo *The unreasonable effectiveness of mathematics in the natural sciences* con estas palabras: "el milagro de la adecuación del lenguaje de la matemática para la formulación de las leyes físicas es un don maravilloso que ni entendemos ni merecemos. Deberíamos mostrarnos agradecidos por él y esperar que permanecerá siendo válido en la investigación futura y que se extenderá, para bien o para mal, para placer nuestro aunque también tal vez para nuestra perplejidad, a ramas más amplias del saber".

La matemática es también un juego. "Un buen juego matemático vale por muchos teoremas" ha afirmado uno de los grandes matemáticos de nuestro tiempo, J. E. Littlewood. El carácter lúdico de las matemáticas ha estado presente en ellas desde antiguo. Como ya hemos tenido ocasión de ver, Aristóteles considera las matemáticas hasta cierto punto como la ocupación de los ociosos sacerdotes egipcios. El tipo de reto intelectual que la geometría elemental o la teoría de números presenta difiere muy poco del que se encuentra en multitud de acertijos populares. Una buena proporción de la matemática seria de nuestros textos debe su nacimiento a un profundo acertijo o juego matemático. Así comenzó, por ejemplo, la topología: "el problema que, según entiendo es muy bien conocido, se enuncia así: en la ciudad de Königsberg, en Prusia, hay una isla, llamada Kneiphof, rodeada por los dos brazos del río Pregel. Hay siete puentes A, B, C, D, E, F y G, que cruzan los dos brazos del río [*véase la figura 4a*]. La cuestión consiste en determinar si una persona puede realizar un paseo de tal modo que cruce cada uno de los puentes una sola vez. Se me ha informado de que mientras unos negaban la posibilidad de hacerlo y otros lo dudaban, nadie sostenía que fuese posible realmente". Con estas palabras introduce Euler el problema de los siete puentes de Königsberg, iniciando así un nuevo género de consideraciones matemáticas que darían lugar a la moderna topología. El problema que se propone aquí es equivalente, como se observa fácilmente reduciendo la isla y las zonas de tierra a sendos puntos y los puentes a arcos que los unen, al de dibujar la figura indicada sin levantar el lápiz del papel y sin recorrer cada línea más de una vez. Un problema semejante al de dibujar, con las mismas condiciones, las figuras indicadas a continuación que todos hemos conocido de niños.

Ha habido muchos problemas famosos con aspecto de enigmas de la esfinge que han dado lugar a la creación de teorías matemáticas. Tal tipo de problemas abunda extraordinariamente en el campo de la teoría de números. Muchos están aún por resolver. He aquí algunos: ¿Es cierto que para ningún entero n mayor que 2 y ningún trío de números enteros positivos x, y, z mayores que 0 se verifica $x^n + y^n = z^n$? (problema de Fermat); ¿es cierto que todo número par es suma de dos números primos? (conjetura de Goldbach); ¿existen infinitos pares de números primos (p_1, p_2) tales que $p_2 = p_1 + 2$? (problema de los números primos hermanos).

El famoso problema de los cuatro colores (¿es posible colorear adecuadamente un mapa cualquiera con cuatro colores o existirá un mapa que necesite cinco colores para que los países vecinos se distingan apropiadamente?) nació como curiosidad de un estudiante de Cambridge en 1852. El problema ha estimulado por más de un siglo el trabajo de innumerables matemáticos y ha sido resuelto por Appel y Haken, de la Universidad de Illinois, en 1976 [*véase* "La solución del problema del mapa de cuatro colores", por Kenneth Appel y Wolfgang Haken; INVESTIGACIÓN Y CIENCIA; diciembre 1977]. Las implicaciones del método de resolución de esta cuestión, que nació como un rompecabezas curioso, sobre la estructura de la matemática futura están aún por ver. Lo que ahora es el teorema de los cuatro colores (cuatro colores bastan) ha sido demostrado mediante la intervención, hasta ahora totalmente imprescindible, del computador. Es decir, no se conoce ninguna cadena de razonamientos del tipo de la matemática de siempre, asequible por la mente humana sin la ayuda de aparatos externos, que conduzca al teorema. Este hecho nuevo en la historia de la matemática plantea cuestiones profundas y muy interesantes sobre el porvenir de la matemática. ¿Será tal vez imposible construir una demostración manejable sin computador? ¿Qué grado de certeza proporciona una demostración de este tipo? Nuestra evidencia es sustituida por la fiabilidad del computador. ¿Cuántos y qué tipo de problemas abiertos en matemáticas esperan una solución semejante? ¿Acaso el problema de Fermat o la conjetura de Goldbach puedan resolverse de modo análogo? ¿Modificará sustancialmente el computador nuestra forma de proceder como matemáticos?

El juego matemático sigue atrayendo intensamente tanto a matemáticos como a no matemáticos. Prueba de ello

es, por ejemplo, la constancia del éxito de la sección permanente de Martin Gardner por tantos años en esta revista. De vez en cuando aparecen en el mundo juegos matemáticos, en nuestros días el cubo de Rubik, que por su capacidad de atracción se convierten en el entretenimiento intenso de millones de personas por muchas horas. Es una verdadera lástima que el contenido lúdico y artístico de la matemática no sea explotado más a fondo por los enseñantes a todos los niveles, pero sobre todo en los primeros pasos de la educación matemática. A mi parecer la incapacidad confesada para las matemáticas de muchas personas altamente cultivadas, competentes en otras actividades y profundamente inteligentes, se debe sobre todo a un bloqueo psicológico inicial, originado por los métodos equivocados de enseñanza.

Vayamos, por fin, a la matemática como aventura del pensamiento. Heidegger ha hablado de "eterno aburrimiento del avance rectilíneo de la matemática", observación compartida por muchos que se maravillan de que en esta ciencia haya todavía algo por hacer. Tal ignorancia puede ser disculpada por el hecho de que así como para hablar de otras muchas actividades del espíritu humano sólo se precisa un poco de audacia y otro poco de información, de las matemáticas sólo se puede hablar desde dentro; sólo quien ha asimilado el método matemático y ha tratado de manejarlo está capacitado para entender lo que la creación matemática comporta.

La matemática es una actividad profundamente humana y como tal ha participado y participa intensamente de las vicisitudes históricas del hombre. Está sujeta, al igual que todas las artes, a las modas, corrientes y catástrofes que afectan al hombre. Por ejemplo, la decadencia del pueblo griego y la supremacía del romano, menos dotado para la especulación que el primero, fue causa de un estancamiento de la matemática de Occidente por muchos siglos. La adopción de un simbolismo inapropiado retrasó en unos cincuenta años el desarrollo de las matemáticas en Inglaterra. A una escala temporal más reducida, es evidente que existen modas y corrientes que producen fluctuaciones más o menos intensas en la actividad matemática a todos los niveles, influyendo tanto en la formación de quienes comienzan como en la actividad de los matemáticos de vanguardia.

La matemática participa efectivamente de modo intenso de los avatares históricos del hombre. Pero es que, además, en su mismo progreso intrínseco, la matemática tiene todos los trazos de una apasionante aventura del espíritu. En ella abundan los momentos de incertidumbre y confusión, tales como el apercibimiento por los pitagóricos de la existencia del número irracional, como su nombre indica (ἄλογος ἀριθμός), un monstruo que no encaja en el armazón cristalino y límpido de la matemática pitagórica.

En lugar de enfrentarse con el reto que representaba este contratiempo, o de saber convivir con él, la matemática griega abandonó un tanto el desarrollo del número y se dedicó más intensamente al cultivo de la geometría, lo que tuvo consecuencias importantes en la evolución de la matemática. La aparición del número imaginario fue encajada con más pragmatismo por los algebristas del Renacimiento, que aprendieron a manejarlo aceptablemente incluso sin entender su naturaleza. Ha habido otros muchos períodos importantes de perplejidad y polémica intensa. Las paradojas de Zenón en el siglo v (Aquiles y la tortuga, la dicotomía, la flecha,...), que tratan de poner de manifiesto la incapacidad humana para dominar los procesos infinitos, debieron causar una sensación de callejón sin salida semejante a la que los matemáticos del siglo xx han experimentado cuando empezaron a aparecer las paradojas de la teoría de conjuntos que cuarteaban lo que había comenzado a ser considerado como base de todo el edificio de las matemáticas. Ante esta situación y ante los teoremas descubiertos por Gödel y los lógicos modernos, que ponen bien en claro, entre otras cosas, la incapacidad de la matemática para colocar sus propias bases fuera de toda duda (imposibilidad de demostrar la consistencia, es decir, la ausencia de contradicción, del edificio matemático), constituye expresión típica del sentir de los matemáticos contemporáneos las siguientes palabras de N. Bourbaki en *L'architecture des mathématiques*: "Creemos que la matemática está destinada a sobrevivir y que jamás tendrá lugar el derrumbamiento de este edificio majestuoso por el hecho de una contradicción puesta de manifiesto repentinamente, pero no pretendemos que esta opinión se base sobre otra cosa que la experiencia. Es poco, dirán algunos. Pero desde hace 25 siglos los matemáticos tienen el hábito de corregir sus errores y de ver así su ciencia enriquecida, no empobrecida. Esto les da el derecho de arrostrar el porvenir con serenidad".

Taller y laboratorio

¿Se ha detenido a pensar por qué, y cómo, hierve el agua?

Jearl Walker

El sonido del agua en ebullición preparada para el té o café me resulta tan familiar que apenas si reparo en él. Pero siempre podría decirles cuándo está el agua en plena ebullición, lista para verter en la cafetera o en la tetera. Otra cosa es dar una interpretación pormenorizada de lo que acontece dentro del recipiente. Pese a la dificultad que entraña, se conocen ya casi todos los aspectos termodinámicos del asunto, y así sabemos que ocurren más cosas que las que puedan patentizarse a la vista o al oído.

A la hora de abordar el fenómeno me inspiré en la obra *A Heat Transfer Textbook*, de John Lienhard, de la Universidad de Houston. Un complemento muy útil va a ser una gráfica que relaciona el flujo térmico (o cantidad de calor por unidad de superficie que pasa al agua por unidad de tiempo) con la diferencia de temperatura entre la cacerola (concretamente el fondo de ésta) y la masa de agua. El fondo de la cacerola se pone a más temperatura que el agua, en cuyo seno se establece, consecuentemente, un régimen de convección y se inicia la evaporación.

Creía yo que tal gráfica mostraría una simple curva creciente, en la que se vería que la cantidad de energía cedida al agua aumentaba con la temperatura de la cacerola. No es así: la curva alcanza un máximo, para decaer rápidamente y volver a aumentar a continuación. En las temperaturas más bajas el flujo térmico crece, con la temperatura, pero, conseguido el máximo, desciende bruscamente.

Las temperaturas del agua y del fondo de la cacerola suelen medirse tomando como punto de comparación la temperatura de ebullición normal del agua para la presión existente. El resultado se conoce como temperatura de saturación. Del agua a la temperatura de ebullición se dice que está saturada. Cuando el agua se calienta por encima de su temperatura de saturación, sin cambiar de fase de líquido a gas, se dice que está recalentada. En la gráfica adjunta [*véase la figura 2*], en la escala de temperaturas se reseña la diferencia entre las temperaturas del fondo del recipiente y la de saturación del agua.

Cabría suponer que, en las etapas iniciales de calentamiento, la capa de agua situada por encima mismo de la superficie caliente alcanzaría enseguida la temperatura de saturación, en cuyo caso formaría burbujas de vapor. Pero no: la convección mantiene la temperatura del agua por debajo de la de saturación. ¿Qué sucede? El calor cedido al agua rebaja la densidad de ésta, obligándola a ascender a causa del empuje de Arquímedes. Y el agua más fría desciende al fondo, donde se calienta.

De ese modo, la convección transporta a la superficie el calor transferido al agua. En la superficie, el agua puede cambiar de estado y, si el flujo térmico es suficientemente bajo, este tipo de transferencia prosigue hasta que todo el agua se torna vapor, desprendiéndose de la superficie.

Si el flujo térmico es mayor, el proceso de ebullición alcanza pronto la etapa de ebullición nucleada. Entonces, empiezan a formarse burbujas en el fondo de la cacerola. Si la fuente calorífica es más potente que los hornillos de cocina, el agua atraviesa otra etapa: se forman burbujas de vapor en numerosos lugares y se fusionan en columnas, chorros y volutas de vapor que se elevan hasta la superficie y escapan. En ambas etapas de ebullición nucleada la formación de burbujas de vapor incrementa la transmisión de calor desde la superficie calentada.

Si se hierve agua dulce en el hornillo, en la primera etapa de ebullición nucleada aparecen dos tipos de burbujas: las de vapor y las formadas por el aire disuelto en el agua. Las primeras son burbujas propiamente dichas; las segundas se llaman también pompas. Conforme el agua recibe calor, el aire se ve obligado a salir de la solución formando burbujas aisladas en el fondo.

Distinguir entre ambos tipos de burbujas es fácil atendiendo a su comportamiento. Las pompas (aire) permanecen estables sobre el fondo hasta que se desprenden y suben. Las burbujas (vapor), que no son nítidas, se forman y se rompen para volver a constituirse a razón de 30 a 60 veces por segundo. Estas, aun cuando se desprenden del fondo y comiencen a elevarse, se rompen dentro del agua más fría que encuentran más arriba. (Esta etapa recibe el nombre de ebullición subenfriada, porque sólo está saturada o recalentada la capa de agua más baja.) El agua que haya sido calentada previamente, en un hornillo o la del circuito de agua caliente de un edificio, no genera pompas porque de ella ya se ha eliminado la mayor parte del aire.

En la primera etapa de ebullición nucleada comienzan a aparecer también burbujas de vapor aisladas en ciertos puntos del fondo, que son los llamados focos de nucleación. En éstos se generan burbujas en gran número, pero todas se rompen al poco de formarse o cuando comienzan a elevarse por el agua más fría de encima. Si el flujo térmico prosigue sin interrupción, el agua situada hasta una buena altura por encima del fondo caliente se calienta lo suficiente para que las burbujas lleguen más lejos. Y, si pueden alcanzar la superficie libre, revientan y sueltan el vapor.

Gran parte de la vaporización en las burbujas tiene lugar durante la ascensión por el líquido recalentado. Podría creerse que la mayor parte del calor que transmite el fondo de la cacerola se invierte en formar allí burbujas. No es cierto. La mayor parte del mismo va a parar al agua en contacto con él, bien en las zonas situadas entre los focos de formación de burbujas, bien en el reflujo de agua que se organiza en esos puntos cuando se rompe o se desprende una burbuja.

Cuando crece la temperatura durante la ebullición nucleada, crece nota-

blemente el número de focos de nucleación en el fondo de la cacerola. Entonces las burbujas se fusionan casi al instante en columnas, chorros y volutas de vapor que se quedan en las inmediaciones de los puntos de formación. En esta fase de ebullición, el calor se transfiere al agua con gran rapidez. Las burbujas nacen y abandonan el fondo a una velocidad tal que el agua fluye a borbollones hacia el fondo para recibir allí calor.

La disminución rápida de flujo térmico se presenta a una temperatura algo más elevada: durante la etapa de ebullición llamada transitoria. Se trata de una situación muy peligrosa para los sistemas de enfriamiento industrial; en efecto, mientras en el receptáculo sigue entrando calor, el cedido al agua desciende bruscamente. El aumento de temperatura puede dañar el receptáculo, por cuya razón el régimen transitorio se llama también de ebullición crítica y punto de quemado.

Durante la ebullición crítica todo aumento de la temperatura del recipiente se traduce en una disminución de la cantidad de calor que el agua arrastra consigo. La causa de ello estriba en el vapor que se genera sobre el fondo, ya que la transmisión del calor a través del vapor es más lenta que a través del agua. Cuanto mayor sea la temperatura del recipiente, tanto más se dilatará el vapor y más deficiente será la transmisión de calor.

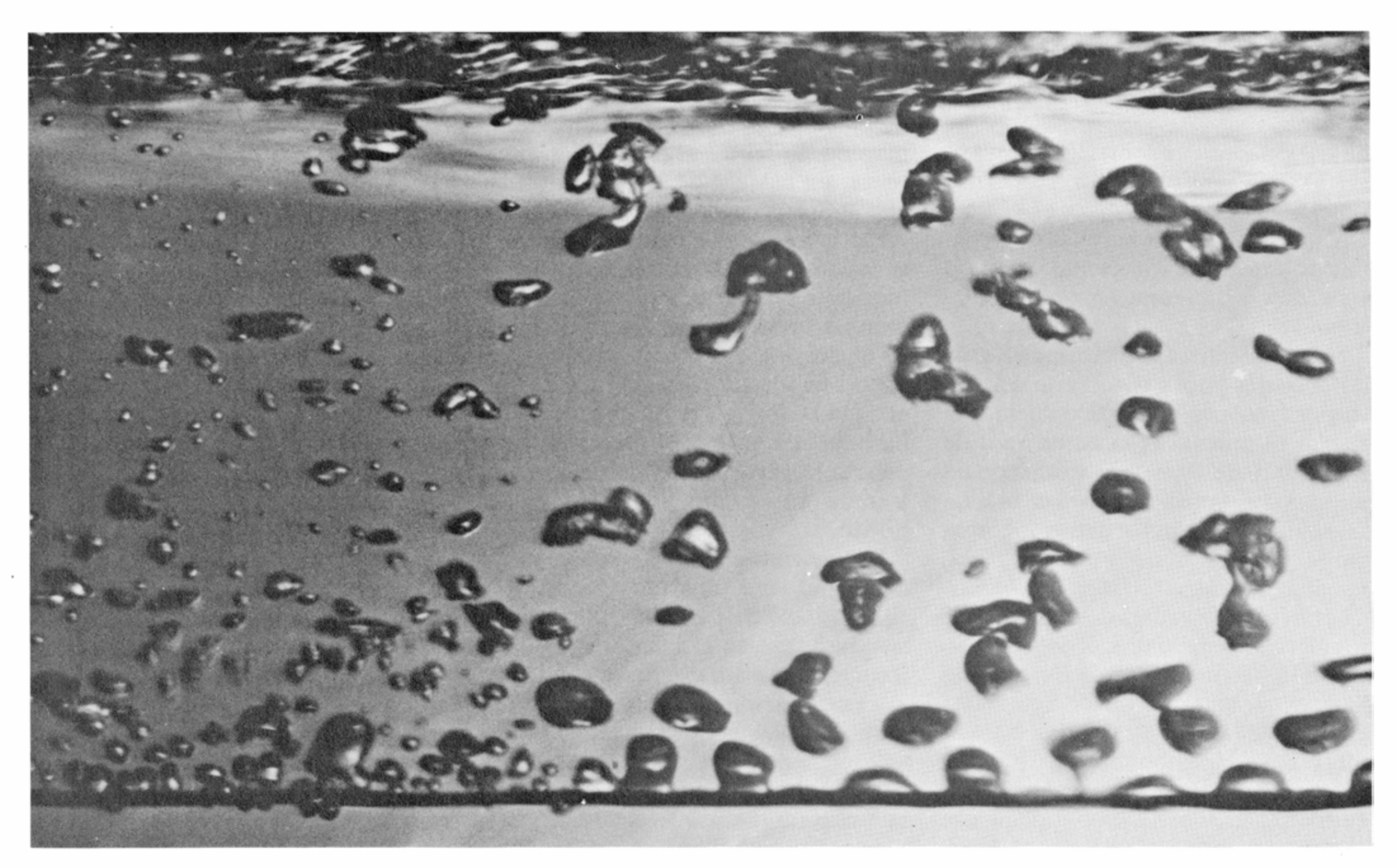
En la etapa de ebullición a muy alta temperatura, el flujo térmico vuelve a aumentar. En esta etapa, denominada de ebullición laminar, sobre el fondo del recipiente se extiende una capa de vapor uniforme. Y, aunque el flujo térmico empiece siendo escaso por culpa del vapor, la alta temperatura acaba impulsando el calor a través de la capa.

El agua que se haga hervir en un hornillo de cocina no alcanza, empero, temperaturas tan elevadas. Lo que no impide que se manifiesten todas las características de la ebullición hasta el régimen de burbujas aisladas. Lienhard sugiere los experimentos siguientes. Póngase una cacerola con agua del grifo sobre un fuego de cocina encendido al máximo. Pasado un rato, se observará parte de la circulación que resulta cuando el agua caliente abandona el fondo de la cacerola y sube hacia la superficie. Esa circulación se hace visible merced a las diferencias de densidad que aparecen dentro del agua en virtud del calentamiento, lo cual origina variaciones en el índice de refracción del agua y, por ello, deforma la trayectoria de los rayos que se reflejan en el fondo.

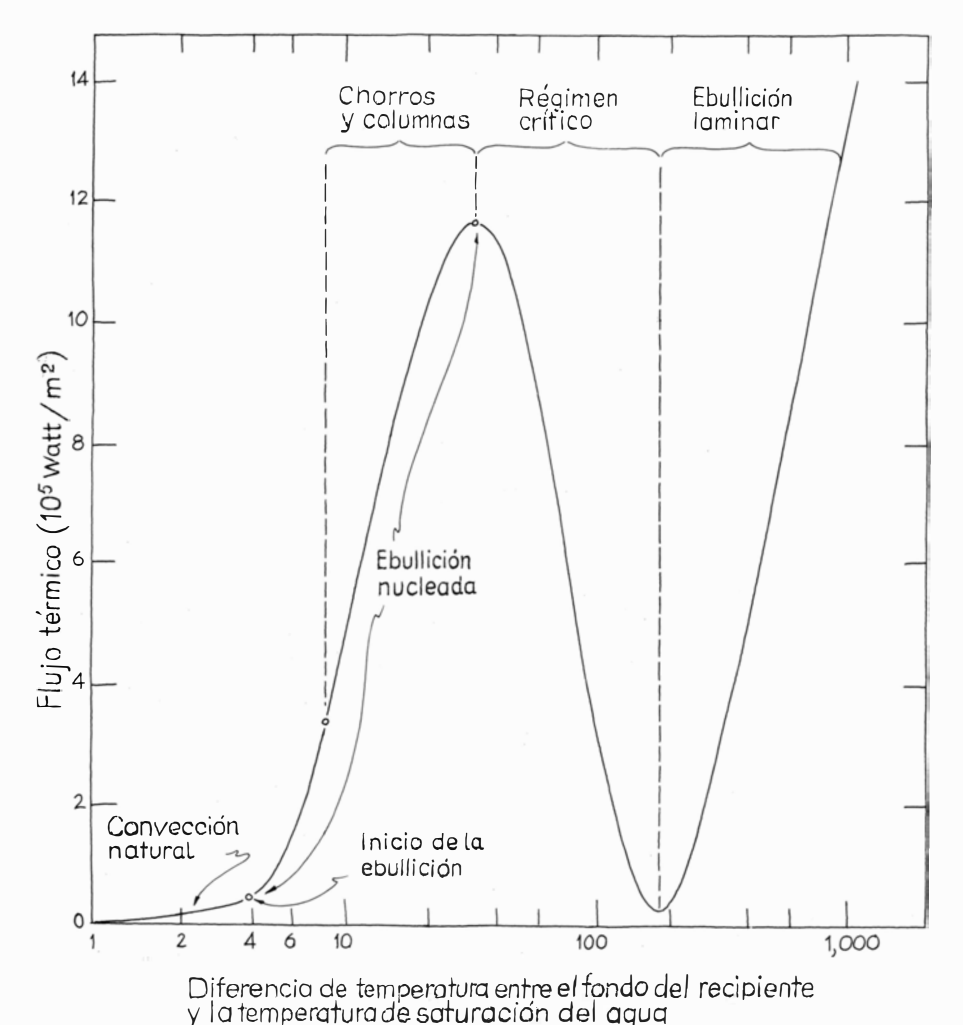
No tardarán en aparecer pompas sobre el fondo, posiblemente en los lugares más calientes. En mi cocina eléctrica las espiras calientan de fuera adentro, lo que me faculta para saber cuándo comienzan a calentar las espiras interiores con sólo observar la distribución de las burbujas de aire. Con el propósito de lograr un calentamiento más uniforme interpongo un trozo de aluminio (de 25 milímetros de grueso y un poco más ancho que la cacerola) entre las resistencias y la cacerola. Serviría mejor un trozo más grueso.

Poco después de aparecer las pompas, la cacerola arranca a "cantar". Este sonido revela la rápida creación, oscilación y rompimiento de burbujas de vapor en el fondo. Cuando cada una de ellas se rompe, emite un silbido o clic agudos. (Las pompas no lo hacen porque no se rompen.) A la vez que el agua prosigue calentándose, son más las burbujas de vapor que emiten su nota de ruptura y la cacerola produce todo un murmullo. Mi experiencia me dice que el agua no se halla todavía suficientemente caliente para hacer café.

El ruido se desvanece paulatinamente conforme crece el número de bur-



1. Fotografía de John H. Lienhard de dos etapas de la ebullición de la acetona alrededor de un alambre caliente: **a la izquierda**, *ebullición nucleada y,* **a la derecha**, *ebullición laminar*



2. Las tres etapas de la ebullición

bujas de vapor que sobreviven a su travesía por el agua más fría y alcanzan la superficie, donde revientan en pequeñas salpicaduras. El burbujeo que entonces sobreviene es la señal de que el agua está lista, o sea, en plena ebullición.

Según Lienhard, el movimiento de las burbujas puede observarse más estrechamente con un estroboscopio. Para ello se calienta agua en un vaso de laboratorio de Pyrex. (El vidrio común podría saltar en añicos con el calor.) La frecuencia de destello del estroboscopio se ajusta entre seis y diez hertz. Para que la luz del estroboscopio se difunda convenientemente, delante de éste se coloca un trozo de papel de seda o de vidrio deslustrado. Con las luces de la habitación apagadas puede retardarse el movimiento de las burbujas sobre el fondo del recipiente.

La generación de burbujas se inicia en los minúsculos hoyitos y pequeñas grietas existentes en la base de todo recipiente. Aunque el fondo de cualquier cacerola parezca liso, en realidad está plagado de tales irregularidades diminutas, en algunas de las cuales queda aire atrapado cuando en el recipiente se vierte agua por primera vez; sin que en su interior pueda penetrar agua,pues lo impide la tensión superficial. Estas diminutas bolsas de aire hacen de focos de nucleación de burbujas de aire y de vapor, que luego crecen a medida que el agua recibe calor. En el caso de las burbujas de vapor, el crecimiento se debe a la vaporización del agua recalentada en el aire contenido en el hoyito.

Teóricamente, el radio de la burbuja lo establece una situación de equilibrio. En la superficie de la burbuja, la presión de vapor en el interior de ésta se iguala a la presión exterior ejercida por el agua circundante más la tensión superficial del agua existente en la superficie de separación entre el agua y la burbuja. Lo habitual, sin embargo, es que esta última sea inestable frente a toda perturbación que acorte su radio, pues entonces la tensión superficial se hace incontenible para la presión de vapor interna y la burbuja se rompe. Pero, si por cualquier causa, aumenta el radio de la burbuja, el líquido situado en la superficie de separación se vaporiza en el interior de la misma, con lo cual crece.

El radio teórico de una burbuja estable depende en parte de la temperatura del agua circundante. Al principio, el agua no se recalienta mucho. Entonces, el radio estable es grande y, por tanto, no suelen desarrollarse las burbujas, ya que la mayoría de los focos de nucleación están ocupados por burbujas diminutas. Conforme aumenta la temperatura del agua, disminuye el valor del radio estable y las burbujas tienen más posibilidad de formarse y crecer.

Al subir la temperatura del agua, lo hace también el flujo térmico sobre el agua, por lo menos en la etapa de burbujas aisladas. Pero el flujo térmico entrante no va directamente a las burbujas, sino que éstas actúan a modo de bombas minúsculas que continuamente impulsan al agua más fría a ponerse en contacto con la superficie caliente de la cacerola.

Supongamos que el agua no esté suficientemente caliente para que una burbuja se escape hacia la superficie del agua. Esta burbuja, ubicada sobre un foco de nucleación, dispone de un radio superior al estable y crece, por tanto, mientras el agua se vaporiza en su interior. Pero ese crecimiento fuerza a que la superficie propiamente dicha penetre en una zona de agua más fría y, así, en esa parte de la burbuja, se condensa vapor, aunque siga vaporizándose el agua en la superficie caliente. Con ello, la parte superior de la burbuja comienza a romperse y, seguidamente, se deshace la burbuja entera. La ruptura de la burbuja provoca una minúscula circulación de agua: el agua recalentada del fondo se proyecta hacia arriba y la situada en la parte superior desciende. Y así se calienta el agua más fría. Este mecanismo aumenta la rapidez de transmisión de calor al agua.

Más avanzado el proceso de calentamiento las burbujas de vapor se desprenden de la superficie caliente. Aquí ocurre que la burbuja crece en su foco de nucleación absorbiendo calor del líquido recalentado circundante. Cuando escapa, en su estela arrastra agua más fría, agua situada más lejos de la superficie caliente. Esta circulación acrecienta el flujo térmico hacia el agua, que se va acrecentando, además, porque en esta etapa posterior de ebullición nucleada entran en actividad más focos de nucleación.

En la práctica, la velocidad a la que se transmite calor al agua durante la ebullición nucleada depende parcialmente de la naturaleza de la superficie caliente. El radio de toda burbuja que pretenda formarse en un foco de nucleación está determinado hasta cierto punto por el tamaño del hoyito origen de la nucleación. Los hoyitos grandes y los provistos de fondo plano son, no obstante, inoperantes respecto a la nucleación de burbujas. Cuanto menor sea el hoyito, menor tenderá a ser la burbuja que se forme sobre él. Ahora bien, los hoyitos más pequeños quizá no incidan en las primeras fases de la generación de burbujas, pues los radios de las que nazcan en ellos son menores que el de estabilidad y las burbujas propenderán a romperse. Mas, a medida que crece la temperatura del agua contigua a la superficie caliente, el radio de burbuja estable disminuye y los hoyitos adquieren mayor importancia.

En general, el flujo térmico para una temperatura dada es mayor en una cacerola cuyo fondo presente una superficie rugosa que en una cacerola en la que dicha superficie sea lisa y pulimentada. Esto lo he comprobado sirviéndome de una de mis cacerolas de acero, que calenté a través de una placa de aluminio colocada sobre un hornillo eléctrico. El fondo de esa cacerola estaba surcada de arañazos, consecuencia de años de servicio y de limpiezas con estropajo. Al calentar agua fría, procedente del grifo, en el perímetro de la cacerola se apiñaron las pompas, evidentemente porque en esa parte de la superficie exterior del fondo era mejor el contacto con el alumnio.

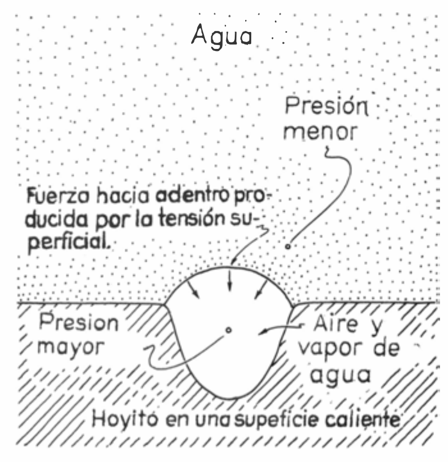
Vi en la superficie interior una nítida red que formaban los arañazos acumulados a través de los años. Las burbujas de aire no se distribuían reflejando tal dibujo. Se presentaban dispersas más o menos uniformemente, al igual que las burbujas de vapor. Innegablemente, los arañazos visibles resultaban de un tamaño excesivo para tomar parte en la formación de burbujas.

A la vez que la cacerola se iba calentando más, la formación de burbujas se extendía desde fuera hacia el centro. Sin embargo, cuando las burbujas de vapor cercanas al perímetro comenzaron a aparecer de golpe, la actividad en las proximidades del centro decreció para acabar desapareciendo. Y eso, ¿por qué? Con toda seguridad, la pieza de aluminio estaba entonces más caliente que nunca. Mi opinión es que tal hecho, aparentemente curioso, se debe a la rapidez del flujo térmico por el anillo perimétrico. En efecto, la nucleación de las burbujas se inicia en ese anillo y, tan pronto las burbujas de vapor tienen la posibilidad de desprenderse, el flujo térmico en el mismo es mucho mayor que el flujo térmico sobre la parte central. O sea, el anillo absorbe calor del aluminio interpuesto en cantidad suficiente para impedir que el agua situada junto al centro de la cacerola se recaliente lo bastante para que en esa zona prosiga la nucleación de burbujas.

Tras vaciar la cacerola y recalentarla brevemente para que se secara el interior, raspé en tres zonas contiguas al perímetro. En una de ellas froté con suavidad dos veces con un estropajo de acero. En otra, hice dos arañazos fuertes con la hoja de un destornillador de acero. Raí un arañazo doble con la punta afilada de una varilla de aluminio. Luego volví a llenar la cacerola con agua del grifo fría y la puse al calor.

En la zona frotada con el estropajo, la generación de burbujas resultó igual que en las zonas del perímetro que no habían tocado. Los surcos abiertos por el destornillador parecían asimismo inoperantes en la nucleación de burbujas y, aunque éstas aparecían en la vecindad, no guardaban relación alguna con los surcos. Los arañazos más someros practicados por la varilla de aluminio generaban grandes cantidades de burbujas, primero de aire (pompas) y luego de vapor.

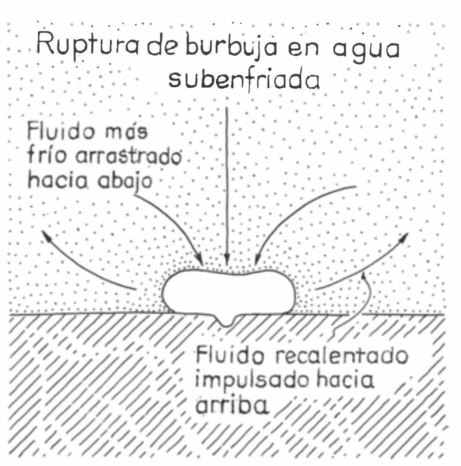
Cuando el agua entró en ebullición turbulenta me fue imposible observarla adecuadamente por lo que, para reducir la ebullición, de tanto en tanto izaba la cacerola. Entonces, aunque la generación de burbujas cesaba casi al instante en todos los demás puntos, en los arañazos practicados con la varilla de aluminio proseguía durante 20 o 30 segundos. No me parece que esos arañazos fuesen realmente focos de nucleación, habida cuenta de su excesiva longitud; lo más probable es que la varilla de aluminio rebrincase un poco al surcar el metal y vaciase pequeños hoyitos en los arañazos, los cuales sirvieron luego de focos de nucleación.

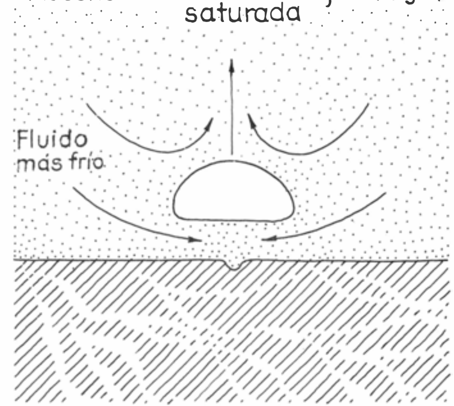


3. Presión interna y externa sobre una burbuja

Podemos construir un modelo que describa el crecimiento de una burbuja en una irregularidad superficial, imaginando un hoyito de forma cónica que se hunda en la superficie calefactora. Inicialmente, el aire encerrado en él estará cubierto de una superficie de agua curvada hacia arriba y se encontrará a la misma temperatura que el agua circundante. La presión en el interior de una tal bolsa de gas superará la presión hidrostática circundante, ya que debe contrarrestar la fuerza dirigida hacia dentro que viene producida por la tensión superficial en la superficie de separación entre el agua y el gas. El radio de curvatura de esta superficie de separación lo determina el cociente entre la tensión superficial del agua y la diferencia entre las dos presiones.

El crecimiento de la burbuja se inicia cuando llega a recalentarse la capa de





4. Patrón de circulación a raíz de la ruptura y elevación de una burbuja

agua contigua a la superficie caliente. Entonces, el agua que se halla en la superficie de separación se evapora en el interior del hoyito, aumentando de volumen. Con ello aumenta el radio de curvatura y disminuye la fuerza dirigida hacia dentro debida a la tensión superficial. (Algo parecido acontece cuando hinchamos un globo esférico con la boca. Al principio, el radio de curvatura del globo es pequeño y las fuerzas dirigidas hacia dentro, contrarias a la dilatación, son grandes. Pero cuando el radio de curvatura se hace mayor, tales fuerzas se atenúan y cuesta menos seguir inflando el globo.)

Al hacerse menor la fuerza que se ejerce sobre la bolsa de gas, la diferencia de presiones entre el interior de ésta y el agua circundante se reduce. Esta situación persiste mientras la evaporación hacia el interior de la bolsa va elevando la superficie de separación hasta el borde del hoyito. Llegado este momento, la evaporación subsiguiente y la dilatación de la bolsa empujarán la superficie de separación aún más hacia arriba, pero la burbuja no abandonará todavía el borde.

Seguidamente, la burbuja llega a un punto crítico. Su expansión hacia arriba reduce el radio de curvatura de la superficie de separación, incrementándose la fuerza hacia dentro debida a la tensión superficial. En el punto crítico, el radio de curvatura de la superficie de separación tiene su valor mínimo y, en correspondencia, la intensidad de la fuerza debida a la tensión superficial será máxima. En tal estado, la diferencia entre las presiones en el interior de la burbuja y en el líquido circundante es la máxima.

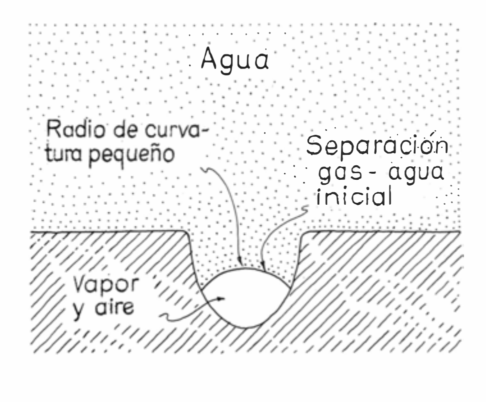
En esas circunstancias pueden ocurrir dos cosas. Si el agua circundante se encuentra suficientemente recalentada, en la superficie de separación continuará la evaporación. Entonces la burbuja crecerá de tamaño mientras la superficie de separación se aparta del borde del hoyito. Así aumentará el radio de curvatura y decrecerá la fuerza hacia dentro inducida por la tensión superficial, en virtud de lo cual la burbuja seguirá creciendo y quizá acabe separándose de la superficie caliente, dejando tras de sí una bolsa de vapor que servirá de núcleo de formación de la burbuja siguiente.

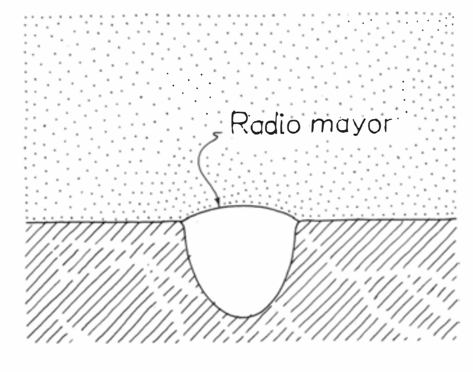
Pero cabe también que el agua circundante no esté suficientemente recalentada para proveer de vapor a la burbuja suspendida en el borde del hoyito. La burbuja dejará de crecer en ese caso, al menos hasta que la afluencia de calor incremente la temperatura del agua. El punto exacto de calor en que tiene que situarse el agua, para permitir que una burbuja traspase el punto crítico durante su crecimiento, depende, en parte, del tamaño del hoyito. Así, una burbuja en un hoyito relativamente grande (de 10 micrometros de diámetro, digamos) quizá no requiera mucho recalentamiento para atravesar sin dificultad su estado crítico. Y una burbuja en un hoyito pequeño (de un micrometro de diámetro, por ejemplo) llega a su estado crítico con un radio de curvatura menor y, por tanto, con una presión interna mayor. El agua circundante deberá calentarse todavía más para que la burbuja pueda superar el estado crítico. En esto último reside la causa de que los puntos rugosos más pequeños de la superficie caliente tarden más que los grandes en participar en la formación de burbujas.

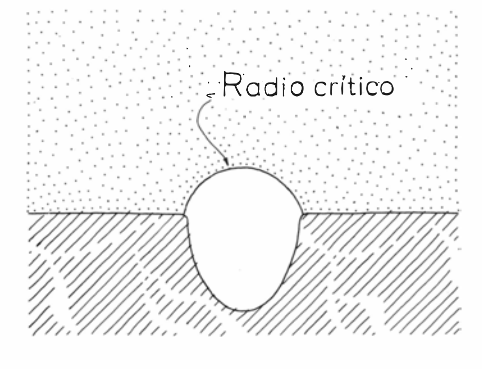
Mi razonamiento, empero, hace abstracción de que muchísimos puntos de rugosidad no sean cónicos. Como también prescinde de que haya hoyitos que no atrapen aire. Cuando se vierte agua en una cacerola, hay un manto de líquido que avanza sobre los hoyitos y, entonces, que el agua llene o no por completo un hoyito depende de dos factores: la geometría de la cavidad y el ángulo de contacto entre el manto de agua y el metal. Este ángulo es función de la tensión superficial del agua en la superficie de separación entre el aire y el metal. Si es mayor que el ángulo de conicidad del hoyito, el agua en su avance llegará al otro lado del hoyuelo antes que al fondo de éste y, de ese modo, quedará aire atrapado que servirá de foco de nucleación. Si no es así, el aire será desalojado del hoyito por el avance del agua y no podrá servir de foco de nucleación.

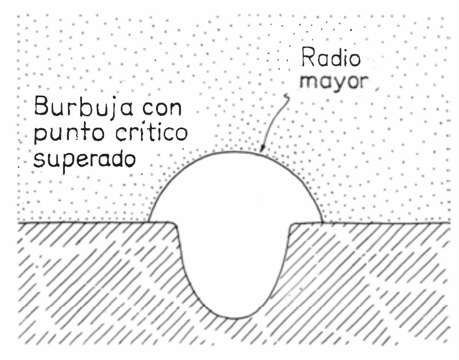
En un hornillo de cocina, resulta imposible llegar a la etapa de la ebullición que marca el paso de la formación de volutas y columnas a la fase de quemado del régimen crítico. (Pero vale la pena detenerse en la física subyacente.) ¿Cómo es posible que la velocidad a la que se transmite el calor cambie, de un modo bastante brusco, desde su valor máximo a otro tan bajo que ponga en peligro el recipiente? Hay que tener en cuenta que la velocidad de transmisión máxima se alcanza exclusivamente a causa de la formación de burbujas sobre la superficie caliente. Entonces, cuando las burbujas se sueltan y ascienden, provocan una agitación continua en el seno del agua que lleva agua más fría sobre la superficie caliente. Conforme el proceso de ebullición se acerca a la etapa de máxima transmisión de calor, sobre la superficie caliente surge una red de burbujas, que se desprenden rápidamente y ascienden en columnas o volutas, mientras el agua más fría se ve arrastrada hacia el fondo.

Las columnas de vapor ascendente forman como una malla encima de la









5. Etapas consecutivas del crecimiento de una burbuja

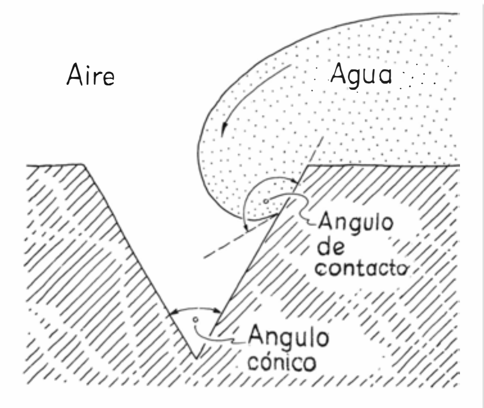
superficie caliente. La situación en la que un fluido menos denso (una capa de vapor) está debajo de otro más denso (el agua) se conoce como inestabilidad de Rayleigh-Taylor. Aquí las columnas de vapor ascendentes se ven forzadas a adoptar una configuración aproximadamente geométrica. En esta etapa, el calor sigue transmitiéndose eficazmente a través del agua, pero a medida que la superficie del fondo se calienta más y más las columnas de vapor deben subir más deprisa al objeto de que se mantenga el régimen de ebullición.

Al final, se mueven con tal rapidez que sus superficies verticales se hunden produciendo ondas en lo que se conoce como inestabilidad de Helmholtz. Como consecuencia, la transmisión pierde eficacia y el proceso de ebullición llega rápidamente al punto de quemado, mientras que la afluencia de calor a la superficie del fondo rebasa la capacidad del sistema para transportar vapor hacia la supeficie libre.

No es el agua, desde luego, el único líquido que puede calentarse o romper a hervir en una cocina. Muchas veces he pasado por la desalentadora experiencia de calentar más de la cuenta alguna salsa espesa. Con las prisas, me ocurre que doy demasiado calor. Luego deduzco que, puesto que la superficie sigue estando fría, el fondo debe estarlo también. Pero, repentinamente, unas enormes burbujas barbotean en la superficie y salpican de salsa todo el hornillo.

Jeffrey C. May, de la Cambridge School de Weston, Massachusetts, tuvo la amabilidad de llamarme la atención sobre tan común error culinario. Atribuye mi falta a la despreocupación por la viscosidad del fluido. Cuando se calienta agua, en la etapa inicial la convección del fluido consigue transportar calor desde el fondo hasta la superficie, donde el agua puede evaporarse. Pero en una salsa espesa la convección no tiene efectos prácticos y el fluido se recalienta en el fondo hasta que se generan burbujas de gran tamaño, aunque la superficie esté todavía fría. Las burbujas se separan del fondo bruscamente y suben hasta la superficie, donde revientan con estrépito.

A la hora de cocer pasta sigo dos reglas para hervir el agua. La primera es calentar un volumen de agua aproximadamente triple del volumen que ocupa la pasta. ¿Por qué? Con menos agua, la ebullición nucleada se detendría al introducir la pasta en la cazuela y la transmisión de calor se limitaría a convección. O sea, la parte superior del agua se enfriaría, dejando la pasta en agua tibia y complicándole el cronometraje al cocinero. Pero si la cazuela contiene la cantidad de agua correcta, la adición de la pasta afectará apenas a la ebullición.



6. Inundación de un hoyito cónico

Jamás he estado muy seguro acerca de la lógica de mi segunda regla. Todas las recetas que he leído mandan agregarle sal a la pasta. Si es para sazonarla, como se supone, ¿por qué no debe añadirse a la salsa? En las explicaciones se afirma que la sal eleva la temperatura de ebullición del agua y que así la pasta se cuece antes. Algo de verdad hay en el aserto, puesto que todo solvente (agua, aquí) que contenga un soluto no volátil (sal, aquí) tiene un punto de ebullición más alto que en estado puro.

En una ebullición turbulenta la presión de vapor del agua debe igualar a la presión atmosférica sobre el agua. Cuando ésta no está saturada, la presión de vapor es inferior a la atmosférica. Si al agua se agrega un soluto como sal, se reduce la presión de vapor en el agua. La causa estriba en que la presencia de la sal (iones de sodio y de cloro) en la superficie frena la posibilidad de que las moléculas de agua escapen. Cuando la solución de agua y sal alcanzan la temperatura de ebullición del agua pura, la presión de vapor es aún inferior a la que se ejercía en ausencia de sal. Pero la temperatura aumenta más, hasta que la presión de vapor termina por igualar la atmosférica; sólo entonces la solución romperá a hervir.

¿Acelera la adición de sal la cocción de la pasta? Añadí media cucharadita de sal de una sola vez a algo menos de dos litros de agua hirviendo. Aunque dupliqué la cantidad de dos cucharaditas que indica una receta clásica, la temperatura del agua subió algo más de un grado. Hay que deducir que la sal sirve sólo de condimento.

Libros

Farmacología, corrosión de los metales y Noam Chomsky

Francisco G. Valdecasas, Pedro Molera y Juan José Acero

MODERN PHARMACOLOGY, por Charles R. Craig y Robert E. Stitzel. Editorial Little, Brown and Co. Boston, 1982. El libro está concebido, según expresan los propios autores, como un libro de texto "organizado y estructurado específicamente para llenar las necesidades de los estudiantes y de los médicos de hospitales universitarios". Se trata, pues, de una obra para la enseñanza de la farmacología y texto de consulta rápido para el médico clínico. Parece una cosa fácil, pero no lo es. En la enseñanza de la farmacología concurren una serie de circunstancias, en parte consecuencia de su propia historia, en parte derivadas de la propia naturaleza de su contenido, que suponen graves dificultades a la hora de su aprendizaje.

Históricamente, la farmacología es una ciencia reciente, con poco más de un siglo de existencia. Por ello, todavía no ha calado hondo en los estudios médicos, y los estudiantes que, antes de entrar en las facultades, ya están deslumbrados por ciencias como la anatomía, con un pasado glorioso y nombres legendarios, o con la fisiología, o con la patología, apenas han oído hablar de la farmacología cuyo nombre les suena más bien a brebaje, a pócima o a untura con términos enrevesados que, además, no son los que se usarán cuando lleguen a enfrentarse con la práctica profesional. Una vez completados los estudios de la anatomía macroscópica, a finales del siglo pasado y aún bien avanzado el presente, la gran pasión de la medicina era la morfología microscópica normal y patológica que, junto con la patología, permitieron los grandes avances diagnósticos y la estructuración de las enfermedades sobre base científica. Comprendida la enfermedad, sus causas y su evolución, la vieja terapéutica, basada en el desequilibrio de los humores, quedaba irremediablemente destinada al desván de las cosas inútiles. Los grandes maestros de la medicina la menospreciaban y se hizo famosa la contestación de uno de ellos al interno de turno que, en la mano el libro de prescripciones, preguntaba qué pondría en él tras escuchar una brillante disertación sobre la enfermedad y los síntomas que aquejaban al paciente –"lo que Vd. quiera, ¡qué más da!".

Lo más importante en un libro docente es el conseguir despertar el interés del lector, la pasión por conocer su contenido. A pesar de los éxitos alcanzados por la farmacología en los últimos, pongamos, cincuenta años, en que ha sido el factor decisivo para lograr la victoria sobre tantas enfermedades, contribuyendo, además, al conocimiento de la fisiología a nivel molecular, el estudiante, en general (hablo del estudiante español que es el que conozco mejor, pero no me parecen muy distintos los de otros países), no se acerca a ella ilusionadamente. Piensa que es demasiada química. Que su terminología es demasiado complicada y que, por último, todas esas dificultades se verán aliviadas a la hora de la verdad, es decir, frente al enfermo, por la propaganda de las casas productoras de fármacos que, con nombres distintos y con machacona reiteración, le harán aprenderse aún sin quererlo, con su propaganda, especialidades e indicaciones.

Es evidente que todo ello es, en parte, verdad. Pero también es evidente el riesgo que se corre con esta actitud, que podríamos llamar de "espíritu de curandero". ¿De qué sirve haber afinado todo lo posible en un diagnóstico, conociendo etiología, patogenia y pronóstico, si no aprovechamos todos esos finos detalles a la hora del tratamiento? ¿Para qué gastar tiempo y esfuerzo, medios refinados y análisis costosos, si a la hora del tratamiento resolvemos el problema con lo que hemos oído del "visitador médico"?

Puede pensarse que exagero, quizá por deformación profesional. Pero es que, en el campo de la terapéutica, el médico está asaltado por multitud de "demonios familiares". Detengámonos por un momento en la complicación terminológica. En los estudios y en la literatura científica, el estudiante se enfrenta con una nomenclatura totalmente distinta de la que luego se le ofrece como medio terapéutico. Cada agente terapéutico, es decir, cada fármaco, está en el mercado, no con otro nombre, sino con decenas de nombres distintos. ¿Hemos de aprender toda la nomenclatura? Imposible. No queda otro remedio que acudir a libros clave que, como es sabido, van generalizándose. Pero éstos, muy útiles en sí mismos, no pueden sustituir a un conocimiento básico de la acción farmacológica y de su mecanismo, a pesar de que tengan entrada por indicaciones y exposición sinóptica de efectos.

El "qué más da" del viejo maestro ya no es válido, ni siquiera a la hora de decidir por una u otra sustancia de idéntica o parecida acción farmacológica. Los medicamentos actuales son eficaces, no puede dudarse de ello, pero por eso mismo tienen su riesgo. Uno de los *desiderata* de la farmacología es, precisamente, disminuir el riesgo terapéutico. Pero disminuir no es anular. El riesgo no será siempre el mismo en todos los enfermos. Será mayor o menor, dependiendo de muchas circunstancias; unas congénitas (variabilidad individual), otras adquiridas, ya por procesos patológicos concomitantes, ya por la ingestión anterior o simultánea de otros medicamentos. Lo mismo podemos decir de la propia acción terapéutica. Sólo un conocimiento científico de la sustancia farmacológica, en sus interacciones con el organismo, nos dará la base para un tratamiento óptimo en cada caso concreto.

Este conocimiento científico de los medicamentos, necesario para un ejercicio profesional de la medicina, sólo puede adquirirse a través del estudio en un libro de texto moderno. Con razón, los autores de *Modern Pharmacology* han destacado este aspecto tan necesario. Porque la farmacología ha evolucionado con tal rapidez que ya no se trata de incluir más y más sustancias medicamentosas en los textos de enseñanza (cosa que, evidentemente, hay que hacer), sino de incorporar los nuevos conceptos que permiten explicar el

modo de acción de las drogas. Por ejemplo la teoría de los "receptores".

La interacción fármaco-sustancia biológica no se ha comprendido hasta el desarrollo de esta teoría, que en la actualidad podemos decir que ya ha alcanzado la realidad de los hechos probados. Se tiene a P. Ehrlich como el primer autor que propuso esta teoría, al sugerir la imagen de las "cadenas laterales". Sin embargo, el concepto de receptor actual es cosa mucho más compleja y mejor definida. Además, es un concepto que abarca no sólo la interacción fármaco-célula, sino también la interacción hormona-célula o célula-célula, hasta el punto de que en la actualidad no podemos imaginar una acción biológica que no tenga en su base la interacción sustancia química-receptor, y esto ya se trate de procesos inmunitarios, de procesos de regulación hormonal o de regulación nerviosa. La teoría de los receptores, trabajosamente elaborada desde los primeros momentos para explicar la afinidad selectiva de los quimioterápicos, pudo explicar más tarde la acción de los transmisores nerviosos o servir de base a la farmacología cuantitativa. Hoy en día representa el punto de unión entre la acción químico-biológica y la teoría informática aplicada a las relaciones de las células en los organismos o de los organismos en las sociedades.

Es por ello necesario y despierta el interés del lector un capítulo suficientemente extenso de farmacología general (un 10 por ciento del total del libro). *Modern Pharmacology* está concebida a la manera frecuente en los textos americanos, aunando la aportación de diversos autores especialistas en los correspondientes apartados. Un total de 47 han sido llamados a este fin. Tal número de colaboradores parece ventajoso para asegurar la competencia y la "modernidad" de la redacción y garantiza que el texto esté "a la última". Pero también debe reconocerse que semejante elenco de redactores conlleva una gran dificultad para mantener la necesaria unidad que un libro de texto exige, acoplando la extensión de los distintos capítulos a la importancia real de los mismos en armonía con la totalidad del texto. No hay duda de que ésta, que es la labor fundamental de los "editores" (en el sentido inglés de la palabra), se ha conseguido a la perfección.

Otros capítulos de la farmacología moderna, tales como el de la cinética de los fármacos, la metabolización y eliminación, la excreción y biodisponibilidad, etcétera, reciben la debida extensión. La repercusión de todos estos parámetros en las distintas facetas de la medicina práctica, dosificación, pauta y ritmo de los tratamientos, duración de los mismos y aparición de los fenómenos de acumulación y toxicidad crónica se estudian igualmente en su justa amplitud.

En la parte descriptiva del libro debemos destacar su extensión. Esta es intermedia entre la de los grandes tratados, que sólo se pueden considerar como libros de consulta, y los resúmenes sinópticos propios para los que muestren el mínimo interés por el tema. En la literatura anglosajona de la materia farmacológica, un libro de esta extensión se echaba en falta. Otra cosa a señalar de la parte descriptiva es la orientación predominantemente terapéutica tendente a servir los problemas que se pueden suscitar en el ejercicio de la clínica. Así pues, muchas sustancias farmacológicas se citan en diversos capítulos, incluso con el detalle necesario, según su indicación. Pongamos como ejemplo a los corticoides, tratados, según es debido, en el capítulo de la farmacología hormonal, pero también en otro capítulo dedicado al control del asma, problema de obvia importancia práctica.

Terminaremos diciendo que los "editores" pueden sentirse satisfechos de haber conseguido su propósito de hacer un texto "no tan detallado que abrume con la exposición de datos a los estudiantes, ni tan breve y simple que les prive de las bases fundamentales para comprender en su plenitud la acción terapéutica de las drogas", según sus propias palabras. Añadamos nosotros un elogio a la excelente presentación y a la ilustración suficiente, pero no excesiva, de fórmulas químicas y esquemas ilustrativos. (F. G. V.)

CORROSIÓN Y PROTECCIÓN DE LOS METALES EN LA ATMÓSFERA, por Sebastián Feliu y Manuel Morcillo. Ediciones Bellaterra, S. A., Barcelona 1982. Un volumen 24 × 17 de 246 páginas, con numerosas figuras y tablas. Encuadernado en rústica. Los metales más ampliamente utilizados por el hombre –hierro, aluminio, cobre, níquel y todas sus aleaciones–, se alteran y pierden sus adecuadas propiedades, mecánicas, físicas y químicas, cuando, sin protección alguna, permanecen en contacto con la atmósfera debido a un proceso más o menos lento denominado *corrosión*. Los metales citados se encuentran en la naturaleza en estado combinado, fundamentalmente formando óxidos; los procedimientos industriales de obtención metálica tienen por objeto transformar los compuestos en metales, pero por razones termodinámicas inexorablemente se vuelven a convertir en compuestos. El objetivo de las técnicas de protección metálica no es otro que dificultar la viabilidad de la exigencia termodinámica, preservando las piezas metálicas acabadas, lo que suele conseguirse tan sólo en parte.

Las pérdidas materiales globales de corrosión en España se estiman en 75.000 millones de pesetas anuales, según datos del año 1973, de los cuales un elevado tanto por ciento corresponde a la corrosión en ambientes naturales (atmósfera, agua y suelo). Más del 50 por ciento de las pérdidas materiales causadas por la corrosión lo fueron por corrosión atmosférica. Parece razonable postular que el ambiente origina en España pérdidas anuales por corrosión del orden de los 30.000-40.000 millones de pesetas. Según un informe del Ministerio de Tecnología del Reino Unido, se podría ahorrar cerca de un 25 por ciento de los costes de corrosión con un mejor conocimiento de los mecanismos y fundamentos de los fenómenos corrosivos y la apropiada aplicación de las técnicas de protección actualmente asequibles.

Se proponen los autores hacer una revisión crítica de las investigaciones realizadas y de los conocimientos actuales, tratando de condensarlos de forma asequible, asimilable y útil para el lector. Este objetivo lo consiguen con creces, avalados por la sólida, profunda y acreditada experiencia que ambos poseen en los campos de la corrosión y de la protección de los metales. Sebastián Feliu es profesor de investigación y vicedirector del Centro Nacional de Investigaciones Metalúrgicas (CSIC) y Manuel Morcillo es investigador del citado centro.

El libro está dividido en 14 capítulos. A modo de introducción, en el primero se describen los mecanismos y las variables de la corrosión atmosférica. La acción de la humedad y de la temperatura atmosféricas, junto con otros datos meteorológicos, se estudian en el capítulo 2, para en el siguiente presentar una recopilación de esos datos referentes a España, haciendo hincapié en la estimación del tiempo de humectación. En los capítulos 4 y 5 se examina, pormenorizadamente, la influencia de los contaminantes atmosféricos en los procesos corrosivos.

Dado el enorme interés del acero en construcciones expuestas a la atmósfera, se detiene el capítulo sexto en una serie de relaciones entre los datos de corrosión del acero y las variables me-

teorológicas y de contaminación y el tiempo de exposición. Seguidamente (capítulo 7) se compara la resistencia a la corrosión de los aceros "patinables" con la del acero dulce. Los capítulos 8, 9 y 10 abordan los metales no férreos típicos: el cinc, el cobre y el aluminio. El onceavo se dedica al análisis de los datos recogidos en la exposición de muestras de acero, cinc, cobre y aluminio en una limitada serie de estaciones de ensayo situadas en España en puntos de distinta meteorología y diverso grado de contaminación atmosférica. La influencia de los recubrimientos protectores en la corrosión se enfoca de modo resuelto y generalizado en los capítulos 12 y 13, donde se estudian los materiales metálicos resistentes a la corrosión, los recubrimientos metálicos, las pinturas, etcétera. El libro termina, con el capítulo 14, estudiando los ensayos de corrosión atmosférica, la evaluación de los factores meteorológicos y contaminantes y la calibración de la agresividad corrosiva atmosférica. Valorando cuantitativamente los factores ambientales se identifica, con rigor y garantía, el tipo de atmósfera corrosiva.

Se trata de un libro de temática muy interesante, escrito con rigor científico, conciso en la exposición, ilustrado con ponderación y editado con esmero. Está destinado no sólo al investigador y experto en el campo de la corrosión y de la protección, sino también a cualquier técnico encargado del proyecto y mantenimiento de las construcciones metálicas o preocupado por su integridad y duración en la atmósfera. (P. M.)

SOBRE NOAM CHOMSKY: ENSAYOS CRÍTICOS, por Gilbert Harman y otros, recopilación e introducción de Gilbert Harman, traducción de Violeta Demonte y Juan Carlos Moreno (si bien Soledad Valera consta como traductora del ensayo de Jerrold Katz). Madrid, Alianza Editorial, 1981, 354 páginas. Esta obra es una recopilación de ensayos cuyo elemento en común estriba en versar sobre cuestiones suscitadas por, o emparentadas con, la obra del lingüista norteamericano Noam Chomsky. La casi totalidad de estos ensayos tienen un valor incuestionable, pero a juicio de quien esto escribe hay que destacar las contribuciones de Robert Lees ("Recensión de *Estructuras sintácticas*") –publicada en 1957 y a la que el propio Chomsky ha atribuido un papel decisivo en el reconocimiento que la citada obra llegó a alcanzar–, John Ross ("Extractos de *Restricciones sobre las variables en sintaxis*"), de filósofos como Willard Van Quine ("Reflexiones metodológicas sobre la teoría lingüística actual"), Hilary Putnam ("Algunos problemas de la teoría de la gramática"), David Lewis ("Lenguas, lenguaje y gramática") y Donal Davidson ("Una semántica para las lenguas naturales"). Estos tres últimos, conocidos bien en el mundo de habla castellana por los especialistas tan sólo, se hallan entre lo más destacado de la filosofía de habla inglesa, por lo que hay que celebrar que sus escritos empiecen a ser difundidos entre nosotros, aunque sea de esta forma tan fragmentaria e incidental. Dicho esto, hay que señalar que, de todos los ensayos de la recopilación, sólo el de John Searle ("La revolución chomskyana en la lingüística") había sido ya traducido a nuestra lengua y publicado con un título ligeramente distinto.

Los ensayos se agrupan en cuatro apartados: los hay que se ocupan de problemas de teoría lingüística, los que apuntan a diversas cuestiones metodológicas de esta disciplina, los que atienden a temas filosóficos –justificando el recurso a este rótulo en razones puramente académicas– y los que versan sobre problemas de psicolingüística. En justicia, sin embargo, más de un ensayo del total podría incluirse en dos o más de estos apartados.

Al primer grupo pertenecen las contribuciones de Ross, ya mencionada, la segunda parte de la recensión de *Language and Mind*, otra de las obras de Chomsky, realizada por Gilbert Harman, algunas de las secciones de la recensión de Lees, la contribución de Denis Stampe ("Hacia una gramática de significado") y la de Barbara Partee ("Metateoría lingüística"). Las cuestiones debatidas en este grupo son variopintas: desde una aguda discusión de los principales retos a que se enfrentaba la teoría de las gramáticas generativas a finales de los cincuenta hasta las principales conclusiones a que llegó Ross en su tesis doctoral en cuanto a las restricciones que operan sobre las reglas transformacionales de las gramáticas de la especie mencionada, pasando por una interesante propuesta de Harman sobre el fenómeno de la anáfora pronominal, según la cual los sintagmas nominales de la estructura subyacente deberían concebirse como operadores que ligan variables del tipo de los investigados en la lógica contemporánea. En cuanto a Partee, esta lingüista traza un magnífico resumen de la evolución de la teoría lingüística de corte chomskyano de 1957 hasta principios de la década de los setenta. Cuando esta obra se publicó originalmente en inglés, en 1974, el breve ensayo de Partee tenía una vigencia indiscutible. Hoy nos llega con un lastre, aunque no excesivo, sí inevitable.

Al segundo grupo pertenecerían los ensayos de Putnam, Quine, Thomas Nagel ("Lingüística y epistemología"), parte de las recensiones de Leech y Harman y, en cierto sentido, el de Lewis también. Lees, por ejemplo, presenta una brillante defensa de la tesis de que una gramática no es sino una teoría científica, sorprendiendo en su trabajo su clarividencia para con las implicaciones psicológicas de tales teorías –especialmente en las páginas 91 y 93–, tema éste que saldría a la palestra años después, no sin herir susceptibilidades. De las presuntas causas de estas reacciones se ocupan los ensayos de Quine, Nagel y parte de la recensión de Harman. Los dos primeros tratan de diversos aspectos de la teoría chomskyana de las ideas innatas, y en ellos predomina un obvio talante crítico. Quine aprovecha, además, para volver a descargar sus baterías contra la semántica intensional al uso. Harman opina, aunque no expone las razones que le llevan a pensar así, que polémicas como las que Chomsky y Quine han mantenido se basan de hecho en malentendidos, por lo que es posible, al menos en principio, casar puntos de vista aparentemente tan enfrentados. Putnam se propone, por su parte, elaborar "un entorno conceptual para el tipo de trabajo que Chomsky está llevando a cabo" (pág. 95). Este objetivo era necesario en 1961, fecha en que su ensayo fue publicado. Y entre las consideraciones que Putnam hace en él hay que subrayar la de que ya entonces juzgase posible demostrar que las gramáticas propugnadas por Chomsky generaban cualesquiera conjuntos recursivamente enumerables de oraciones de una lengua. La importancia del tema de la constricción de este tipo de gramáticas para la teoría lingüística quedaba ya apuntado.

Los ensayos filosóficos atienden a cuestiones diversas. Jerrold Katz ("La importancia de la lingüística para la filosofía") se ocupa del tema de la distinción entre forma gramatical y forma lógica, sosteniendo en él que la teoría lingüística puede servir para trazar esta distinción satisfactoriamente. Seguro que Katz tiene razón; pero la defensa de su tesis depende de la vigencia del modelo de análisis gramatical de *Aspectos*, que hace ya años que descansa en paz. Davidson lleva a cabo su conocida propuesta de que una teoría semántica para una lengua natural adopte

la forma de una teoría de las condiciones de verdad de (parte de) las oraciones de dicha lengua. Como casi todos sus ensayos, éste es una pequeña obra de orfebrería, pero –aunque no esencial para su propuesta– Davidson piensa, como Katz, que existe un vínculo directo entre estructura profunda o subyacente (nivel de la descripción lingüística en que se definen las funciones gramaticales de los diversos sintagmas de una oración) y forma lógica (estructura responsable de las propiedades lógicas de una oración). El tiempo, sin embargo, parece haber ido en contra de ambos autores en este punto. Lewis atiende al complicado problema de las relaciones entre el análisis sintáctico, el semántico y la dimensión del habla. En este contexto, Lewis formula de manera escueta su importante teoría de las convenciones –es escandaloso que su libro sobre este tema no haya sido traducido todavía a nuestra lengua– y otorga en ella un papel relevante a la teoría pragmática del significado de Paul Grice. El ensayo de Stampe, antes citado, trata precisamente de ofrecer pruebas de carácter sintáctico en favor de esta misma teoría. Hay que decir, sin embargo, que las reglas transformacionales de Stampe deberían examinarse con mucho cuidado antes de darles la bienvenida a los textos gramaticales, pues transformaciones con un aire tan saludable como las que a primera vista tienen las de Stampe han sufrido posteriormente un sino desgraciado. Es decir, las pruebas puramente sintácticas en favor de su adopción no están nada claras.

Los ensayos de Thomas Bever ("La psicología del lenguaje y las investigaciones estructurales del innatismo") –el único que fue escrito para ser incluido en esta recopilación– y el de este mismo autor, junto con James Lackner y Robert Kirk ("Las estructuras subyacentes de las oraciones son las unidades primarias del procesamiento inmediato del habla") son dos ejemplos bien ilustrativos del género de investigación psicolingüística que se atiene a los postulados mentalistas avalados por Chomsky. Ambos incluyen las conclusiones de trabajo experimental de gran interés y en los dos se apunta la tesis de la realidad psicológica de las estructuras subyacentes de las oraciones, a la par que se matizan y critican algunas de las afirmaciones más extremas de los innatistas de nuestros días. Pese al gran interés de estos ensayos, debemos decir que dentro incluso de las comunidades de lingüistas y de psicólogos afines al generativismo se han alzado voces (como las de Wasow y Steinberg, por ejemplo) que han puesto en tela de juicio la significatividad de investigaciones como aquéllas de las que estos ensayos nos informan.

La recopilación se cierra con una interesante recensión, la que Dell Hymes hizo del libro de John Lyons *Noam Chomsky*. Esta recensión es modélica por muchas razones. Destaca de ella lo relativo al tema de los vínculos entre el Chomsky lingüista y el Chomsky que tan enérgicamente se ha opuesto a la política norteamericana y que tan acerbamente ha criticado a los intelectuales que se despreocupan de las aplicaciones de sus investigaciones y de los fines a los que sirven. Interesante resulta también la crítica de Hymes a algunos de los tópicos con que se dibuja el mapa de la evolución de la lingüística norteamericana hasta la revolución chomskyana. Una evaluación de ésta, a la que Chomsky ha replicado en sus *Reflexiones sobre el lenguaje*, se da en el ensayo de Searle.

En cuanto a la traducción hay que afirmar que en términos generales es impecable, muy acertada sobre todo en cuanto a las decisiones terminológicas adoptadas. Pese a esto, los traductores han sido demasiado condescendientes con ellos mismos al sostener, a propósito del ensayo de Stampe, que 'Juan significó que...' es sólo "algo artificiosa" (pág. 287). No es algo artificioso; lo es del todo. Se entiende que los traductores maticen su decisión diciendo que tomarán expresiones como la anterior en el sentido de 'Juan quiso decir que...'. Ahora bien, aparte de que las tesis de Stampe ya no son verdaderas cuando la discusión hace referencia a estas otras palabras –verdaderas para el caso de la lengua castellana–, esa solución es completamente humptydumtyana.

En lo que toca a la edición de la obra hay que lamentar la existencia de algunas erratas, que no suelen afectar al sentido del texto. La más molesta, quizá, es la de la página 317, en la que la oración (2) debería no confundirse con el párrafo que va tras ella. Lo que sí carece de toda justificación es que se haya decidido no incluir el completo y denso índice de materias que acompaña a los ensayos al final de la obra en su edición en inglés. Adaptarlo a nuestra lengua no hubiese supuesto ningún trabajo hercúleo y habría facilitado en gran medida la consulta y el estudio de una obra como ésta, cuyo solo índice viene a constar de los títulos de los ensayos que contiene. Los editores deberían obligarse a no tomarse licencias como la mencionada. (J. J. A.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

IMPLANTACION DE IONES EN METALES

METASTABLE ALLOY FORMATION. J. M. Poate en *Journal of Vacuum Science and Technology*, vol. 15, págs. 1636-43; 1978.

AMORPHOUS METALS AND ION IMPLANTATION. W. Grant en *Journal of Vacuum Science and Technology*, volumen 15, págs. 1644-49; 1978.

TREATISE OF MATERIALS SCIENCE AND TECHNOLOGY, vol. 18. Academic Press; 1980.

HIGH DENSITY CASCADE EFFECTS. D. A. Thompson en *Radiation Effects*, volumen 56, págs. 105-150; 1981.

SEMIEMPIRICAL STUDY OF METASTABLE ALLOYS OBTAINED BY ION IMPLANTATION IN METALS AND SEMICONDUCTORS. J. M. López y J. A. Alonso en *Physica Status Solidi* (a) vol. 72, págs. 777-781; 1982.

IMAGENES DE RADAR DE LA TIERRA DESDE EL ESPACIO

REMOTE SENSING: PRINCIPLES AND INTERPRETATION. Floyd F. Sabins. W. H. Freeman and Co., 1978.

SPACEBORNE IMAGING RADAR: GEOLOGIC AND OCEANOGRAPHIC APPLICATIONS. Charles Elachi en *Science*, vol. 209, n.º 4461, págs. 1073-1082; 5 de septiembre de 1980.

SEASAT ORBITAL RADAR IMAGERY FOR GEOLOGIC MAPPING: TENNESSEE-KENTUCKY-VIRGINIA. J. P. Ford en *American Association of Petroleum Geologists Bulletin*, vol. 64, n.º 12, páginas 2064-2094; diciembre, 1980.

DESARROLLO DE MAPAS Y BANDAS EN EL CEREBRO

A PHYSIOLOGICAL MECHANISM FOR HEBB'S POSTULATE OF LEARNING. Gunther S. Stent in *Proceedings of the National Academy of Sciences of the United States of America*, vol. 70, número 4, págs. 997-1001; abril, 1973.

A MECHANISM FOR PRODUCING CONTINUOUS NEURAL MAPPINGS: OCULARITY DOMINANCE STRIPES AND ORDERED RETINO-TECTAL PROJECTIONS. Ch. von der Malsburg y W. J. Willshaw en *Experimental Brain Research*; suplemento 1; 1976.

ANATOMY AND PHYSIOLOGY OF EXPERIMENTALLY PRODUCED STRIPED TECTA. M. I. Law y M. Constantine-Paton en *The Journal of Neuroscience*, vol. 1, n.º 7, págs. 741-759; julio, 1981.

ORDENADORES PERSONALES

MICROPROCESADORES. Hoo-min D. Toong en *Investigación y Ciencia*, n.º 14, págs. 80-90; noviembre, 1977.

PERSONAL COMPUTER REPORT. Hoo-min D. Toong y Amar Gupta. Digital Systems Associates, 1982.

MUESTRAS DE VIA LACTEA

AN INTRODUCTION TO NUCLEAR ASTROPHYSICS. Jean Audouze y Sylvie Vanclair. D. Reidel Publishing Co., 1980.

ORIGIN OF COSMIC RAYS: INTERNATIONAL ASTRONOMICAL UNION SYMPOSIUM N.º 94. Dirigido por G. Setti, G. Spada y A. W. Wolfendale. D. Reidel Publishing Co., 1981.

ANOMALOUS ISOTOPIC COMPOSITION OF COSMIC RAYS. S. E. Woosley y Thomas A. Weaver en *The Astrophysical Journal*, vol. 243, n.º 2, 1.ª parte, págs. 651-659; 15 de enero de 1981.

THE ELEMENTAL AND ISOTOPIC COMPOSITION OF GALACTIC COSMIC RAY NUCLEI. R. A. Mewaldt en *Reviews of Geophysics and Space Physics*, en prensa.

INTERACCION ENTRE LA CORRIENTE Y LOS ORGANISMOS SESILES

ANIMALS. R. Riedl en *Marine Ecology*, dirigido por Otto Kinne. John Wiley & Sons, Inc., 1971.

WATER MOVEMENT: PLANTS. H. Schwenke en *Marine Ecology*, dirigido por Otto Kinne. John Wiley & Sons, Inc., 1971.

MECHANICAL DESIGN IN ORGANISMS. S. A. Wainwright, W. D. Biggs, J. D. Curry y J. M. Gosline. John Wiley & Sons, Inc., 1976.

MECHANICAL DIVERSITY OF CONNECTIVE TISSUE OF THE BODY WALL OF SEA ANEMONES. M. A. R. Koehl en *The Journal of Experimental Biology*, vol. 69, págs. 107-125; agosto, 1977.

LIFE IN MOVING FLUIDS. Steven Vogel. Willard Grant Press, 1981.

A LA BUSQUEDA DE NUMEROS PRIMOS

CRYPTOLOGY: THE MATHEMATICS OF SECURE COMMUNICATION. G. J. Simons en *The Mathematical Intelligencer*, vol. 1, n.º 4; 1979.

THE PSEUDOPRIMES TO $25 \cdot 10^9$. Carl Pomerance, J. L. Selfridge y Samuel Wagstaff, Jr., en *Mathematics of Computation*, vol. 35, n.º 151, páginas 1003-1026; julio, 1980.

RECENT DEVELOPMENTS IN PRIMALITY TESTING. Carl Pomerance en *The Mathematical Intelligencer*, vol. 3, n.º 3, págs. 97-105; 1981.

ON DISTINGUISHING PRIME NUMBERS FROM COMPOSITE NUMBERS. L. M. Adleman. C. Pomerance y R. S. Rumely en *Annals of Mathematics*, en prensa.

LAS LUCERNAS DE COSA

COSA I: HISTORY AND TOPOGRAPHY. Frank E. Brown en *Memoirs of the American Academy in Rome*, vol. 20; 1951.

COSA II: THE TEMPLES OF THE ARX. Frank Edward Brown, Emeline Hill Richardson y L. Richardson, Jr., en *Memoirs of the American Academy in Rome*, vol. 26; 1960.

LATE ROMAN POTTERY. J. W. Hayes. British School at Rome, 1972.

COSA: THE MAKING OF A ROMAN TOWN. Frank E. Brown. The University of Michigan Press, 1980.

TEMAS METAMAGICOS

PHILOSOPHY OF MATHEMATICS AND NATURAL SCIENCE. H. Weyl. Princeton University Press; 1949.

SYMMETRY. H. Weyl. Princeton University Press; Princeton, 1952.

GRUNDLAGEN DER MATHEMATIK IN GESCHICHTLICHER ENTWICKLUNG. O. Becker. Alber; Freiburg, 1964.

MAGNITUDES Y LÍMITES DEL PENSAMIENTO MATEMÁTICO. O. Becker. Ediciones Rialp; Madrid, 1966.

MATHEMATIKER ÜBER DIE MATHEMATIK. Dirigido por M. Otte. Springer Verlag; Berlín, 1974.

TALLER Y LABORATORIO

BOILING. G. Leppert y C. C. Pitts en *Advances in Heat Transfer*, vol. 1, dirigido por Thomas F. Irvine, Jr., y James P. Hartnett. Academic Press, 1964.

HEAT TRANSFER IN BOILING AND OTHER PHASE-CHANGE CONFIGURATIONS. John H. Lienhard en *A Heat Transfer Textbook*. Prentice-Hall, 1981.